

<https://dx.doi.org/10.17488/RMIB.46.SI-TAIH.1528>

E-LOCATION ID: e1528

## Detección automática de trastorno de ansiedad generalizada: un enfoque de estrategias de historia de vida

Automatic detection of generalized anxiety disorder: a life history strategy approach.

Ana Luisa Islas Ávila<sup>1</sup> , Alicia Martínez Rebollar<sup>1</sup>  , Ricardo Castrejón Salgado<sup>2</sup> , Laura Ávila Jiménez<sup>2</sup> , Hugo Estrada Esquivel<sup>1</sup> .

<sup>1</sup>Tecnológico Nacional de México/CENIDET, Morelos, México.

<sup>2</sup>Instituto Mexicano del Seguro Social OOAD Morelos - México

### RESUMEN

En este artículo se presenta un modelo de detección del trastorno de ansiedad generalizada (TAG) utilizando algoritmos de aprendizaje automático supervisado. Se analizaron 244 registros obtenidos mediante cuatro cuestionarios que evaluaron aspectos sociodemográficos, estrategias de historia de vida y niveles de ansiedad. A partir de estos datos, se seleccionó un subconjunto de características relevantes, se evaluaron ocho algoritmos supervisados y se optimizaron sus hiperparámetros. El modelo XGBoost mostró el mejor desempeño tanto en modelos sin balancear como con técnicas de balanceo de clases. Los resultados destacan la relevancia de las estrategias de historia de vida como variables predictoras, especialmente en dimensiones como las relaciones interpersonales, la maduración precoz y la impulsividad, las cuales mostraron una asociación significativa con el desarrollo del TAG. Las principales limitaciones del estudio incluyen el tamaño muestral y el enfoque en una población localizada en el centro de México. A pesar de ello, los hallazgos confirman la viabilidad de integrar variables de historia de vida en modelos predictivos de TAG. El modelo XGBoost que implementó la técnica de balanceo *SMOTE-Tomek Links* fue validado y alcanzó una exactitud, sensibilidad, precisión y F1 score de 0.85.

**PALABRAS CLAVE:** aprendizaje automático; clasificación; estrategias de historia de vida; trastorno de ansiedad generalizada.

## ABSTRACT

This article presents a model for detecting generalized anxiety disorder (GAD) using supervised machine learning algorithms. A total of 244 records were analyzed, obtained through four questionnaires that assessed sociodemographic aspects, life history strategies, and anxiety levels. From these data, a subset of relevant features was selected, eight supervised algorithms were evaluated, and their hyperparameters were optimized. The XGBoost model showed the best performance in both unbalanced models and those using class balancing techniques. The results highlight the relevance of life history strategies as predictive variables, particularly in dimensions such as interpersonal relationships, early maturation, and impulsivity, which showed a significant association with the development of GAD. The main limitations of the study include the sample size and the focus on a population located in central Mexico. Nevertheless, the findings confirm the feasibility of integrating life history variables into predictive models of GAD. The XGBoost model implementing the SMOTE-Tomek Links balancing technique was validated and achieved an accuracy, sensitivity, precision, and F1 score of 0.85.

**KEYWORDS:** classification; generalized anxiety disorder; life history strategies; machine learning.

## Autor de correspondencia

DESTINATARIO: ALICIA MARTÍNEZ REBOLLAR

INSTITUCIÓN: TECNOLÓGICO NACIONAL DE MÉXICO/CENIDET,  
MORELOS, MÉXICO.

DIRECCIÓN: TECNOLÓGICO NACIONAL DE MÉXICO/ CENTRO  
NACIONAL DE INVESTIGACIÓN Y DESARROLLO TECNOLÓGICO,  
INTERIOR INTERNADO PALMIRA S/N, COL. PALMIRA, C.P. 62490.  
CUERNAVACA, MORELOS.

CORREO ELECTRÓNICO: alicia.mr@cenidet.tecnm.mx

## Recepción:

16 Febrero 2025

## Aceptación:

21 Abril 2025

## Publicación:

22 Diciembre 2025

## INTRODUCCIÓN

La ansiedad puede experimentarse en diferentes momentos de la vida, sin embargo, el trastorno de ansiedad generalizada se distingue por una preocupación excesiva sobre eventos cotidianos que persiste durante meses y se acompaña de síntomas como tensión muscular, alteraciones del sueño, dificultad para concentrarse e irritabilidad<sup>[1]</sup>. En México, el proyecto Encovid-19 reportó en 2020 una prevalencia de este trastorno entre el 31% y 33% de la población<sup>[2]</sup>. Una manera de identificar factores de riesgo para el desarrollo del trastorno de ansiedad generalizada son las estrategias de historia de vida, un enfoque de la psicopatología evolutiva. En años recientes, los modelos de aprendizaje automático han ganado relevancia en la detección de trastornos mentales, incluyendo la ansiedad, mediante análisis comparativos de técnicas de aprendizaje supervisado<sup>[3]</sup>. Sin embargo, no existe suficiente evidencia sobre modelos que utilicen variables de estrategias de historia de vida. Este enfoque evolutivo permite realizar predicciones sobre la psicología individual y el posible desarrollo de psicopatologías, al considerar una combinación de características conductuales y fisiológicas<sup>[4][5][6]</sup>. Los eventos adversos tempranos predisponen hacia estrategias de vida "rápidas", mientras que los entornos estables fomentan estrategias "lentas" orientadas al bienestar a largo plazo<sup>[7][8]</sup>. Por tal razón, el propósito de esta investigación es desarrollar un modelo de detección automática del trastorno de ansiedad generalizada basado en información relacionada con las estrategias de historia de vida. Este enfoque innovador busca mejorar la comprensión y detección temprana de los trastornos de ansiedad, proporcionando nuevas perspectivas para el diagnóstico clínico. Se implementará una variedad de modelos de aprendizaje automático, incorporando técnicas de selección de características, estrategias de balanceo de clases y optimización de hiperparámetros para garantizar un rendimiento óptimo, confirmado mediante la validación del mejor modelo. Los resultados de esta investigación no solo contribuirán al campo de la salud mental desde una perspectiva técnica, sino que también aportarán conocimientos valiosos sobre los factores que influyen en el desarrollo y manifestación de la ansiedad generalizada.

### Fundamentos teóricos

El aprendizaje automático (*Machine Learning*) es un subconjunto de la inteligencia artificial que se centra en utilizar conjuntos de datos que pueden ser limitados o extensos y algoritmos escalables<sup>[9]</sup>. En el ámbito de la salud mental, se ha utilizado para desarrollar modelos que predicen o detectan trastornos a partir de variables específicas y los algoritmos representativos de aprendizaje supervisado que se utilizan en varios trabajos relacionados a salud mental son Máquinas de Soporte Vectorial, Random Forest, Naive Bayes, K Vecinos más Cercanos, XGBoost, Regresión Logística, AdaBoost y Perceptrón Multicapa<sup>[10][11]</sup>.

La Máquina de Soporte Vectorial (*SVM*, por sus siglas en inglés) es un algoritmo de aprendizaje supervisado utilizado para tareas de clasificación. En su versión lineal, el objetivo principal es encontrar un hiperplano que separe de manera óptima dos clases en un espacio de características, maximizando el margen entre los datos más cercanos de cada clase, denominados vectores de soporte<sup>[12]</sup>. El modelo es definido como:

$$f(x) = w^T x + b \quad (1)$$

donde  $w \in \mathbb{R}^d$  es el vector de pesos y  $b \in \mathbb{R}$  es el sesgo. El objetivo es encontrar los parámetros  $w$  y  $b$  que definen un hiperplano de separación con el máximo margen posible.

*Random Forest* es un algoritmo de aprendizaje supervisado basado en el principio de ensamble de modelos, cuyo

objetivo es mejorar la precisión y la capacidad de generalización mediante la combinación de múltiples árboles de decisión construidos sobre subconjuntos aleatorios de los datos<sup>[13]</sup>. En problemas de clasificación, el modelo construye un conjunto de *Tárboles de decisión*  $\{h_t(x)\}_{t=1}^T$ , cada uno entrenado de forma independiente, y la predicción final para una nueva observación  $x$  se obtiene mediante votación mayoritaria entre los árboles.:

$$\hat{y} = \text{modo}\{h_t(x)\}_{t=1}^T \quad (2)$$

*Naive Bayes* es un clasificador probabilístico basado en el teorema de *Bayes*, que asume independencia entre las variables predictivas dado el valor de la clase<sup>[10]</sup>. El objetivo es estimar la probabilidad posterior  $P(x)$  para predecir la clase más probable de una nueva observación  $x$ :

$$P(x | y) = \prod_{j=1}^d P(x_j | y) \quad (3)$$

*K Vecinos más Cercanos* (*KNN*, por sus siglas en inglés) es un algoritmo basado en instancias que consiste en asignar una clase a una nueva observación  $x$  en función de las clases de sus  $k$  vecinos más cercanos en el espacio de características<sup>[9]</sup>. La distancia entre observaciones se calcula generalmente mediante la distancia Euclídea:

$$\text{dist}(x, x_i) = \sqrt{\sum_{j=1}^d (x_j - x_{ij})^2} \quad (4)$$

*XGBoost* (*Extreme Gradient Boosting*) es un algoritmo de aprendizaje supervisado basado en árboles de decisión donde el principio fundamental está basado en la construcción de un modelo aditivo mediante la combinación secuencial de múltiples modelos débiles, donde cada modelo nuevo corrige los errores cometidos por el conjunto de modelos anteriores<sup>[14]</sup>. Sea  $x_i \in \mathbb{R}^d$  un vector de características y  $y_i$  la variable de salida, el modelo XGBoost busca predecir  $y_i$  como una suma de funciones de predicción  $f_t \in F$ , donde  $F$  representa el espacio de árboles posibles:

$$\hat{y}_i = \sum_{t=1}^T f_t(x_i), f_t \in F \quad (5)$$

La regresión logística multinomial es una extensión de la regresión logística binaria utilizada para modelar problemas de clasificación con múltiples clases mutuamente excluyentes. Es un modelo de aprendizaje supervisado que estima directamente la probabilidad de pertenencia de una observación a cada una de las  $k$  clases posibles, basándose en un conjunto de variables predictoras<sup>[15]</sup>. El modelo define una función discriminante para cada clase  $k$  de la forma:

$$P(y_i = k | x_i) = \frac{\exp(w_k^T x_i + b_k)}{\sum_{j=1}^k \exp(w_j^T x_i + b_j)} \quad (6)$$

donde  $w_k$  es el vector de coeficientes correspondiente a la clase  $k$  y  $b_k$  es el intercepto para la clase  $k$ .

*AdaBoost* (*Adaptive Boosting*) es un algoritmo de aprendizaje supervisado basado en métodos de ensamble secuencial donde el principio fundamental consiste en entrenar de forma iterativa una serie de clasificadores  $h_t(x)$ , cada uno centrado en las muestras más difíciles, y luego combinar sus predicciones mediante una ponderación basada en su precisión<sup>[16]</sup> donde el clasificador final tiene la forma:

$$H(x) = \text{sign}(\sum_{t=1}^T \alpha_t h_t(x)) \quad (7)$$

El Perceptrón Multicapa es una arquitectura de red neuronal artificial de tipo *feedforward* que incorpora una o más capas ocultas con funciones de activación no lineales, lo que le permite modelar relaciones complejas y no lineales entre las variables de entrada y salida<sup>[17]</sup>. Esta red neuronal aprende los parámetros (pesos y sesgos) mediante la minimización de una función de pérdida a través del algoritmo de retropropagación del error, usual-

mente optimizado con métodos de descenso por gradiente<sup>[18]</sup>. Un Perceptrón Multicapa con una sola capa oculta se define mediante las siguientes operaciones:

$$z_i = \phi(W^{(1)}x_i + b^{(1)}) \quad (8)$$

$$y^i = \sigma(W^{(2)}z_i + b^{(2)}) \quad (9)$$

donde  $W^{(1)}$  y  $b^{(1)}$  son los pesos y sesgos de la capa oculta,  $W^{(2)}$  y  $b^{(2)}$  son los pesos y sesgos de la capa de salida,  $\phi$  es la función de activación de la capa oculta y  $\sigma$  es la función de activación de la capa de salida.

El análisis de varianza de una vía (ANOVA) es una prueba paramétrica que compara las medias de tres o más grupos independientes para determinar diferencias estadísticamente significativas<sup>[19]</sup>. Este análisis presupone normalidad en las distribuciones, homogeneidad de varianzas y observaciones independientes. La hipótesis nula sostiene que todas las medias poblacionales son iguales, mientras que la alternativa plantea que al menos una difiere significativamente. Por otro lado, la prueba de *Kruskal-Wallis* constituye una alternativa no paramétrica al ANOVA. Se emplea cuando no se cumple el supuesto de normalidad o cuando los datos son ordinales. Esta prueba evalúa diferencias significativas entre las medianas de tres o más grupos independientes, basándose en los rangos de los datos en lugar de sus valores originales<sup>[20]</sup>. Un valor de  $p < 0.05$  indica que al menos un grupo difiere de los demás.

La prueba de *Friedman* es un método no paramétrico para datos relacionados que permite comparar tres o más condiciones dentro de una misma muestra, especialmente cuando no se cumple el supuesto de normalidad requerido para un ANOVA de medidas repetidas<sup>[21]</sup>. Esta prueba organiza los datos por filas, asigna rangos dentro de cada una y evalúa si las sumas de rangos difieren significativamente entre condiciones. Su hipótesis nula establece la ausencia de diferencias entre las condiciones evaluadas. Cuando la prueba de Friedman revela diferencias significativas, es necesario realizar un análisis post hoc para identificar específicamente qué grupos difieren entre sí. La prueba de *Nemenyi* se utiliza como método no paramétrico para realizar comparaciones múltiples por pares, comparando la diferencia absoluta entre los rangos promedio de cada grupo con un valor crítico establecido<sup>[22]</sup>. Si la diferencia entre rangos supera este valor crítico, se considera que los grupos presentan diferencias significativas.

### Trabajos relacionados

En el contexto de los trastornos de ansiedad, la aplicación de modelos de aprendizaje automático para la detección y predicción ha resultado prometedora. La investigación en este campo ha implementado diversos enfoques, utilizando modelos supervisados tradicionales y combinándolos con técnicas de selección de características, técnicas de balanceo de clases e integrando fuentes de datos mixtas, para optimizar la precisión de las predicciones.

Durante la pandemia de COVID-19, se desarrollaron estudios relevantes sobre ansiedad en poblaciones universitarias. En<sup>[23]</sup> se evaluaron cuatro niveles de ansiedad en 917 estudiantes saudíes, comparando siete clasificadores supervisados. *AdaBoost* y redes neuronales destacaron con AUCs de 0.943 y 0.936, respectivamente, donde la red neuronal obtuvo la mejor exactitud (75%) y F1-score (74.9%), identificando como variables clave el sexo, el apoyo social y el ingreso familiar. Por su parte,<sup>[24]</sup> aplicó *Random Forest* con validación cruzada en una muestra de 2,409 estudiantes, utilizando el GAD-7 y 51 preguntas sobre el impacto del COVID-19. El modelo alcanzó un  $R^2$  de 0.429, evidenciando una capacidad predictiva moderada para identificar síntomas de ansiedad y depresión.

El estudio<sup>[25]</sup> se utilizaron los aspectos sociodemográficos de una muestra de 2,121 estudiantes de Bangladesh para detectar depresión y niveles de ansiedad, donde, de seis algoritmos de aprendizaje automático evaluados, el de Máquina de Soporte Vectorial tuvo el mejor desempeño para la detección de ansiedad con una exactitud del 91.49%, identificándose que el sexo, la edad y las calificaciones semestrales bajas impactan en el desarrollo de la ansiedad. En<sup>[26]</sup> se desarrolló un modelo *CatBoost* para la detección de TAG utilizando características sociodemográficas, de salud, de apoyo social y relacionadas con la percepción sobre COVID-19, obteniendo una exactitud y F1 score de 51.84% e identificando que el estrés percibido diariamente, los cambios en el estado de ánimo, el sexo y patrones de sueño alterados fueron los factores más relevantes para clasificar a los participantes en los diferentes niveles de ansiedad.

Un estudio en Líbano<sup>[27]</sup> utilizó una Máquina de Soporte Vectorial para clasificar los niveles de ansiedad en 329 universitarios, logrando un AUC del 76.01% y una exactitud del 69.70%. En<sup>[28]</sup>, un modelo KNN basado en síntomas de ansiedad clasificó a 284 estudiantes latinoamericanos con una exactitud del 97.83%. El estudio longitudinal de<sup>[29]</sup>, realizado con adolescentes europeos entre los 14 y 23 años, integró variables psicométricas, sociodemográficas, conductuales y de neuroimagen, obteniendo un desempeño moderado (62% de exactitud, AUC de 0.69); factores como extraversion, neuroticismo y volumen caudado bilateral destacaron en el modelo para TAG. Finalmente,<sup>[30]</sup> aplicó XGBoost sobre datos de 4,184 estudiantes franceses, logrando una sensibilidad del 70% y un AUC de 0.67; las variables más influyentes fueron el uso de sustancias, riesgo de hipertensión y esquema de vacunación incompleto.

Diversos estudios han aplicado modelos de aprendizaje automático para detectar ansiedad en distintas poblaciones. En<sup>[31]</sup>, se analizaron señales ECG de 66 estudiantes, logrando una precisión, exactitud y sensibilidad del 80% con un modelo de Bosque Aleatorio. En<sup>[32]</sup>, con datos de 5,186 adultos mayores de la cohorte TUDA, se emplearon modelos supervisados y técnicas de balanceo como SMOTE, alcanzando una precisión del 85.4% con Random Forest. El estudio<sup>[33]</sup> utilizó datos sociodemográficos, conductuales y clínicos de 220 pacientes brasileños para clasificar pensamientos obsesivos, destacando un modelo SVM con 75.2% de sensibilidad, 70.6% de especificidad y un AUC de 0.825. Finalmente,<sup>[34]</sup> analizó una muestra de 111 pacientes psiquiátricos y controles con variables psicométricas y de comportamiento, obteniendo una especificidad del 76.81% y sensibilidad del 69.66% con Bosque Aleatorio.

El resumen general de los trabajos mencionados se encuentra en la Tabla 1.

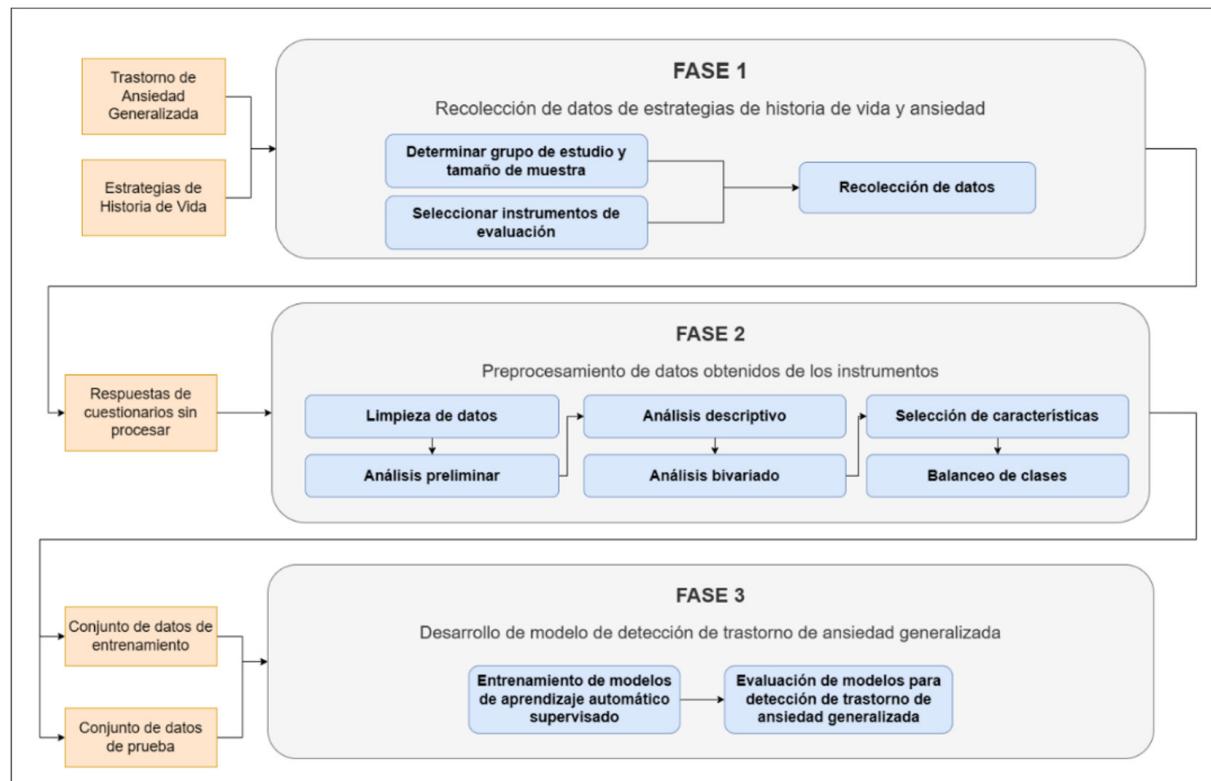
TABLA 1. Descripción general de los trabajos relacionados.

Trabajo	Población	Tipo de datos	Instrumentos de evaluación	Tipo de modelos	Resultados de desempeño
[23]	917 estudiantes de licenciatura de una universidad de Arabia Saudita	Características sociodemográficas y de comportamiento relacionado a COVID-19	GAD-7	KNN, Árboles de decisión, Máquina de Soporte Vectorial, Bosque aleatorio, Red neuronal, Naïve Bayes, AdaBoost	<b>Mejor modelo:</b> Red neuronal Exactitud = 75% Precisión = 76.9% Sensibilidad = 75.4% F1 score = 74.9%
[24]	2409 individuos pertenecientes a una universidad (estudiantes, profesores y administrativos)	Características sociodemográficas y de comportamiento relacionado a COVID-19	PHQ-9, GAD-7	Bosque aleatorio	$R^2 = 0.429$
[25]	2121 estudiantes de licenciatura de universidades de Bangladesh	Características sociodemográficas	PHQ-9, GAD-7	Regresión logística, Bosque aleatorio, Máquina de Soporte Vectorial, LDA, KNN, Naïve Bayes	<b>Mejor modelo:</b> SVM Exactitud = 91.49% Especificidad = 98.53% Sensibilidad = 67.77%
[26]	3878 adultos franceses-canadienses	Características sociodemográficas, de comportamiento, y satisfacción académica	GAD-7, SPS-5, WHOQol-SRPB-BREF	Bosque aleatorio, XGBoost, CatBoost, Regresión lineal, LASSO	<b>Mejor modelo:</b> CatBoost Exactitud = 51.84% F1 score = 51.58% $R^2 = 0.5656$
[27]	329 estudiantes de licenciatura de universidades del Líbano	Características sociodemográficas y de comportamiento	BAI	Regresión logística, Perceptrón multicapa, Máquina de Soporte Vectorial, Bosque aleatorio, XGBoost	<b>Mejor modelo:</b> SVM Exactitud = 69.70% AUC = 0.76
[28]	284 estudiantes de licenciatura de una universidad de Perú	Síntomas auto reportados	GAD-7	KNN, Árboles de decisión, Gradient Boosting, Máquina de Soporte Vectorial, Algoritmos apilados	<b>Mejor modelo:</b> KNN apilado con balanceo de clases Exactitud = 97.83% Precisión = 97.50% Sensibilidad = 98.44% F1 score = 97.88%
[29]	580 adolescentes	Características sociodemográficas, psicométricas, resonancias magnéticas	DAWBA, LEQ	Regresión logística, Máquina de Soporte Vectorial, Bosque aleatorio	<b>Mejor modelo:</b> No se especifica Exactitud balanceada = 62% Especificidad = 71% Sensibilidad = 53% AUC = 0.69
[30]	4,184 estudiantes de licenciatura de una universidad de Francia	Características sociodemográficas, de comportamiento e información clínica	No se especifica	XGBoost, Bosque aleatorio, Máquina de Soporte Vectorial, KNN, Red neuronal	<b>Mejor modelo:</b> XGBoost Sensibilidad = 70% Especificidad = 70% $AUC = 0.67$
[31]	66 estudiantes	Señales cardíacas (ECG)	SAS	Árboles de decisión, Bosque aleatorio, AdaBoost	<b>Mejor modelo:</b> Bosque aleatorio Exactitud = 80% Precisión = 80% Sensibilidad = 80% Especificidad = 73%
[32]	5186 adultos mayores	Características sociodemográficas y de comportamiento e información clínica	HADS	Bosque aleatorio, Máquina de Soporte Vectorial, Gradient Boosting, Perceptrón multicapa, Regresión logística	<b>Mejor modelo:</b> Bosque aleatorio con balanceo SMOTE Exactitud = 85.4% Sensibilidad = 67% Especificidad = 90.3%
[33]	220 pacientes brasileños	Características sociodemográficas, de comportamiento e información clínica	RRS, PSWQ, GAD-7	Máquina de Soporte Vectorial, Bosque aleatorio, Red neuronal	<b>Mejor modelo:</b> SVM Sensibilidad = 75.2% Especificidad = 70.6% AUC = 0.825
[34]	111 pacientes psiquiátricos y de control	Características psicométricas y de comportamiento	STAI-T, RRS, PSWQ	Bosque aleatorio	Especificidad = 76.81% Sensibilidad = 69.66%

**GAD-7**, Escala para el Trastorno de Ansiedad Generalizada-7; **PHQ-9**, Cuestionario de Salud del Paciente-9; **SPS-5**, Escala de Fobia Social-5; **WHOQol-SRPB-BREF**, Cuestionario de calidad de vida de la Organización Mundial de la Salud (BREF): espiritualidad, religión y creencias personales; **BAI**, Inventario de Ansiedad de Beck; **DAWBA**, Evaluación del desarrollo y el bienestar; **LEQ**, Cuestionario de Experiencias Vitales; **SAS**, Escala de Autoevaluación de Ansiedad de Zung; **HADS**, Escala de Ansiedad y Depresión Hospitalaria; **RRS**, Escala de Respuestas Rumiantivas; **PSWQ**, Inventario de Preocupación de Pensilvania; **STAI-T**, Inventario de Ansiedad Estado-Rasgo - Versión Rasgo.

## MATERIALES Y MÉTODOS

En esta sección se detallan los materiales utilizados en este trabajo en el apartado de instrumentos y la metodología implementada (Figura 1) para el desarrollo de un modelo de detección del trastorno de ansiedad generalizada.



**FIGURA 1. Metodología de solución para la detección de trastorno de ansiedad generalizada.**

### Recolección de datos

Se llevó a cabo un estudio con un diseño de encuesta transversal, empleando un enfoque metodológico analítico, retrospectivo y observacional. La población de estudio estuvo conformada por derechohabientes adultos adscritos a la Unidad de Medicina Familiar No. 20 del Instituto Mexicano del Seguro Social (IMSS), ubicada en Cuernavaca, Morelos. Los criterios de inclusión fueron: (1) hombres y mujeres de entre 18 y 59 años que supieran leer y escribir; (2) sin diagnóstico médico de enfermedades neurológicas o neurodegenerativas; y (3) exclusión de personas que hubieran experimentado eventos adversos en el último año. Se aplicaron cuatro instrumentos autoadministrados, con una duración aproximada de 30 minutos por participante. El número total de personas encuestadas fue de  $n=244$ , quienes conformaron la fuente de datos para este estudio.

La Tabla 2 presenta un resumen de las características sociodemográficas, el puntaje del factor K y la categoría de ansiedad de los participantes. La edad promedio fue de 32.04 años (DE = 10.76), con un rango de 18 a 59 años, y un 67% de los participantes fueron mujeres. En cuanto al nivel educativo, el 41% había concluido la educación media superior y el 36% contaba con estudios universitarios completos. La variable principal de análisis fue el nivel de ansiedad, clasificado en tres categorías: sin ansiedad, ansiedad no correspondiente al trastorno de ansiedad generalizada (no TAG) y trastorno de ansiedad generalizada (TAG).

TABLA 2. Resumen de características sociodemográficas, factor K y categoría de ansiedad de los participantes.

Variable	Frecuencia (Porcentaje) n = 244
<b>Sexo</b>	
Mujeres	155 (63.52)
Hombres	89 (36.48)
<b>Edad</b> , en años	
Mediana (p25, p75)	29 (24, 38)
<b>Nivel educativo</b>	
Primaria	6 (2.46)
Secundaria	44 (18.03)
Preparatoria	100 (40.98)
Universidad	87 (35.66)
Posgrado	7 (2.87)
<b>Estado civil</b>	
Soltero/a	108 (44.26)
Casado/a	70 (28.69)
Unión libre	46 (18.85)
Divorciado/a	8 (3.28)
Viudo/a	5 (2.05)
Separado/a	7 (2.05)
<b>Ocupación</b>	
Estudiante	41 (16.80)
Empleado sin estudios profesionales	53 (21.72)
Empleado con estudios profesionales	79 (32.38)
Negocio propio	9 (3.69)
Trabajador independiente	20 (8.20)
Hogar	29 (11.89)
Jubilado/a	4 (1.64)
Desempleado/a	9 (3.69)
<b>Nivel socioeconómico</b>	
Alto	0 (0)
Medio-alto	42 (17.21)
Medio	155 (63.52)
Medio-bajo	43 (17.62)
Bajo	3 (1.23)
Pobreza	1 (0.41)
<b>Personas que habitan en el hogar</b>	
Familia nuclear	132 (54.10)
Padre, madre o ambos	31 (12.70)
Hermanos	9 (3.69)
Pareja/Cónyuge	37 (15.16)
Hijos/as	18 (7.38)
Amigos	3 (1.23)
Solo/a	14 (5.74)
<b>Tiene hijos</b>	
No	119 (48.77)
Sí	125 (51.23)
<b>Hay comorbilidades</b>	
No	183 (75)
Sí	61 (25)
<b>Factor K</b>	
Media (D.E.)	0.53 ± 0.19
<b>Categoría de ansiedad</b>	
Control sano de ansiedad	78 (31.97)
Ansiedad no TAG	92 (37.70)
Trastorno de ansiedad generalizada	74 (30.33)

## Instrumentos

Los instrumentos seleccionados para este estudio permitieron obtener información sociodemográfica de los participantes, aspectos de su estrategia de historia de vida y el nivel de ansiedad en el momento que se aplicaron. En total, se obtuvieron 94 variables de los instrumentos. Para llevar a cabo esta investigación, se aplicaron cuatro instrumentos de evaluación:

1. Cuestionario sociodemográfico. El cual incluyó las variables sociodemográficas de sexo, edad en años, estado civil, nivel educativo, nivel socioeconómico, municipio de residencia, personas con quienes habitan, número de hijos y edad en la que tuvo el primer hijo (si aplicaba), así como la presencia de otras enfermedades crónicas.
2. Cuestionario de estrategias de historia de vida. El instrumento *New Short Form of the Arizona Life History Battery* (K-SF-42) fue desarrollado por Figueredo en 2017<sup>[35]</sup> y está conformado por 42 ítems de autoevaluación diseñados para analizar los distintos aspectos de las estrategias de historia de vida en las subescalas de Perspectiva, Planificación y Control, Calidad de Relaciones Parentales, Apoyo Familiar y de Amigos, Vínculo con Parejas Románticas, Altruismo General y Religiosidad. Los ítems se responden utilizando una escala tipo *Likert* con siete categorías para las primeras 24 preguntas (-3 = *Fuertemente en desacuerdo*, +3 = *Fuertemente de acuerdo*) y cuatro categorías para las preguntas restantes (0 = *Nunca*, 3 = *Mucho*). Su propósito es medir cuantitativamente el factor K, un continuo de estrategia de historia de vida, con altos valores reflejando una estrategia de vida lenta, mientras que los valores bajos indican una estrategia de vida rápida. El factor K se obtiene del promedio de la suma de los puntajes de cada aspecto evaluado y se normalizó para que tuviera un valor entre 0 y 1.
3. Cuestionario de marcadores de estrategias de vida. Las estrategias de historia de vida también se evaluaron utilizando los marcadores del espectro rápido y lento identificados por Del Giudice<sup>[36]</sup>, que incluyen: amabilidad, conciencia, humildad, honestidad, apertura a la experiencia, precocidad sexual, socio sexualidad, apego romántico, orientación a largo plazo, disgusto sexual, impulsividad, propensión a asumir riesgos, búsqueda de sensaciones placenteras, cognición mentalista, habilidad verbal, habilidad visoespacial y maduración sexual. Estos marcadores se midieron a través de 16 preguntas autoadministradas, utilizando una escala *Likert* de cinco categorías (1 = *No me describe en absoluto*, 5 = *Me describe totalmente*).
4. Cuestionario de trastorno de ansiedad generalizada. El Inventory of Preocupación de Pensilvania (PSWQ) fue desarrollado por Meyer en 1990<sup>[37]</sup> y es un instrumento autoadministrable empleado para medir la gravedad de la preocupación crónica, frecuentemente observada en personas con trastorno de ansiedad generalizada. Se responden 16 enunciados relacionados con la preocupación utilizando una escala Likert de 5 puntos que varía desde 1 (*nada característico de mí*) hasta 5 (*muy característico de mí*). Los puntajes posibles están en un rango de 16 a 80, con puntajes más altos indicando mayores niveles de preocupación, mientras que los más bajos señalan menores niveles de esta. De la suma de todos los ítems, se obtienen los siguientes resultados: (1) *Menos de 40 puntos = Controles sanos de ansiedad*, (2) *Entre 40 y 60 puntos = Presencia de ansiedad (no TAG)*, (3) *Más de 60 puntos = Presencia de Trastorno de Ansiedad Generalizada (TAG)*.

Los instrumentos solamente se aplicaron con previa lectura y firma del consentimiento informado por cada participante. En la Tabla 3 se muestran las variables recolectadas de los instrumentos de estrategias de historia de vida

y el Inventario de Preocupación de Pensilvania (PSWQ).

**TABLA 3. Resumen de las variables recolectadas de los instrumentos Marcadores de estrategias de historia de vida, K-SF-42 y PSWQ.**

Marcadores de estrategias de historia de vida	Nuevo formato corto de la Batería de Historia de Vida de Arizona (K-SF-42)	Inventario de Preocupación de Pensilvania (PSWQ)
Amabilidad Consciencia Humildad Honestidad Apertura a la experiencia Experiencias sexuales Socio sexualidad Amor inestable Planes inmediatos Disgusto sexual Decisiones impulsivas Asumir riesgos Cognición mentalista Habilidad verbal Habilidad visoespacial Madurez sexual temprana	Mala situación No me rindo Aprendo situaciones Manera diferente Lado bueno de las cosas Positivo en peores situaciones Apoyo emocional a parientes Bienestar parientes Apoyo emocional conocidos Bienestar amigos Voluntario Contribuyo a causas Soy una persona religiosa Religión es importante Espiritualidad es importante Miembro de una religión Asisto a servicios Qué hacer basado en creencias Pareja no interés Incomodidad privacidad Marcar distancia Fusionarme pareja Demasiada intimidad Pareja no intimidad Atención mamá Esfuerzo mamá Enseñaba mamá Atención papá Esfuerzo papá Enseñaba papá Ayuda familiares Hiciste bien familiares Forma de ser familiares Familia muestra afecto Familia escuchar sentimientos Familia se preocupa por bienestar Ayuda amigos Hiciste bien amigos Forma de ser amigos Amigos muestran afecto Amigos escuchan sentimientos Amigos se preocupan por bienestar Factor K normalizado	Preocupo si no tengo suficiente tiempo Preocupaciones agobian Preocupación por las cosas Muchas circunstancias me preocupan No puedo evitar preocupación Bajo tensión Siempre preocupado/a Difícil dejar preocupaciones Me preocupo al terminar una tarea Todo me preocupa Sigo preocupándome Toda mi vida me he preocupado Preocupación excesiva No puedo detener preocupación Preocupación constante Preocupación por proyectos

### Análisis estadístico

El análisis estadístico es de importancia cuando se desarrollará un modelo de aprendizaje automático ya que permite comprender e interpretar los patrones de datos, lo cual es fundamental para realizar predicciones precisas<sup>[38]</sup>. El análisis estadístico consistió en la limpieza de los datos, un análisis preliminar y el análisis bivariado de las variables sociodemográficas y de estrategias de historia de vida.

El conjunto de datos requirió manejo de valores faltantes y atípicos en varios registros, los cuales, en su mayoría, se originaron durante el proceso de transcripción de información de los cuestionarios físicos a formato digital. Se aplicó una imputación simple basada en la moda, dado que las variables afectadas eran cualitativas. En el análisis preliminar se evaluó la normalidad de las variables cuantitativas mediante histogramas, gráficos de cajas y el esta-

dístico de *Shapiro-Wilk*. La variable edad no presentó normalidad ( $p < 0.05$ ), por lo que se utilizó la prueba no paramétrica de *Kruskal-Wallis* para las categorías de ansiedad. En cambio, la variable factor K cumplió con la normalidad ( $p = 0.14$ ), permitiendo el uso de ANOVA de una vía.

En el análisis bivariado se utilizó el estadístico de Chi Cuadrada y Exacta de *Fisher* para comparar las frecuencias observadas con las frecuencias. Las variables cuantitativas se analizaron con los estadísticos de *Kruskal Wallis* y ANOVA de una vía. Estas variables se contrastaron con la categoría de ansiedad, medida por el puntaje total del PSWQ. El nivel de significancia estadística se estableció en  $p < 0.05$ . Como resultado, las variables que resultaron con relevancia estadística fueron edad ( $p < 0.001$ ), sexo ( $p < 0.01$ ) y factor K ( $p < 0.01$ ).

Se evaluó también la significancia estadística de los 42 ítems que pertenecen al instrumento de K-SF-42 contrastando con las categorías de ansiedad medida por el instrumento PSWQ. Se consideraron también las variables que estuvieran cercanas al nivel de significancia establecido ( $p < 0.05$ ). Se encontraron 19 variables que cumplían con estos criterios, las cuales son: (1) 2 variables de la subescala *Perspectiva, Planificación y Control*, (2) 3 variables de la subescala *Altruismo General*, (3) 5 variables de la subescala *Experiencia en Relaciones Cercanas*, (4) 3 variables de la subescala *Calidad de Relaciones Parentales*, y (5) 6 variables de la subescala *Apoyo Familiar y de Amigos*.

### Selección de características

La selección de características es una actividad fundamental en el aprendizaje automático ya que mejora el desempeño del modelo al identificar y conservar las variables más importantes mientras que descarta las que no aportan información significativa<sup>[19][20]</sup>. Para este proceso se utilizaron las técnicas de Análisis Factorial (AF) y Eliminación Recursiva de Características (RFE).

### Análisis Factorial Exploratorio y Confirmatorio

El Análisis Factorial Exploratorio (AFE) se emplea para examinar los datos y determinar la cantidad de factores cuando no se cuenta con suficiente evidencia previa. Por otro lado, el Análisis Factorial Confirmatorio (AFC) se utiliza para evaluar hipótesis específicas sobre la estructura factorial obtenida del AFE<sup>[41]</sup>. Se realizaron los análisis factoriales, tanto exploratorio como confirmatorio, sobre el cuestionario de marcadores de estrategias de historia de vida. El objetivo fue validar los constructos subyacentes, confirmar la fiabilidad del instrumento y obtener un subconjunto de variables que representaran adecuadamente los marcadores.

Inicialmente, se planteó que el instrumento evaluaba cinco factores relacionados con las estrategias de historia de vida con los 16 ítems: *personalidad, motivación, autorregulación, habilidad cognitiva y maduración sexual*. En el Análisis Factorial Exploratorio (Figura 2), se identificó que el conjunto de ítems X8, X9, X11 y X12 está asociado al factor "Motivación", mientras que los ítems X6 y X16 corresponden al factor "Sexualidad".

Después de obtener los resultados del Análisis Factorial Confirmatorio, se confirmó que los 6 ítems seleccionados evalúan de manera adecuada los factores de "Motivación" y "Sexualidad".

### Eliminación Recursiva de Características (RFE)

La eliminación recursiva de características es un método de selección de características que tiene como propósito eliminar aquellas poco relevantes mientras conserva las más independientes y significativas para mejorar la capa-

ciudad de generalización del modelo<sup>[42]</sup>. Esta técnica emplea un procedimiento iterativo de clasificación de características, basado en la eliminación regresiva. Se implementó en las 19 variables seleccionadas del instrumento K-SF-42, el factor K normalizado, el subconjunto de 6 variables de los marcadores de estrategias de historia de vida y las 2 variables sociodemográficas estadísticamente significativas del análisis bivariado.

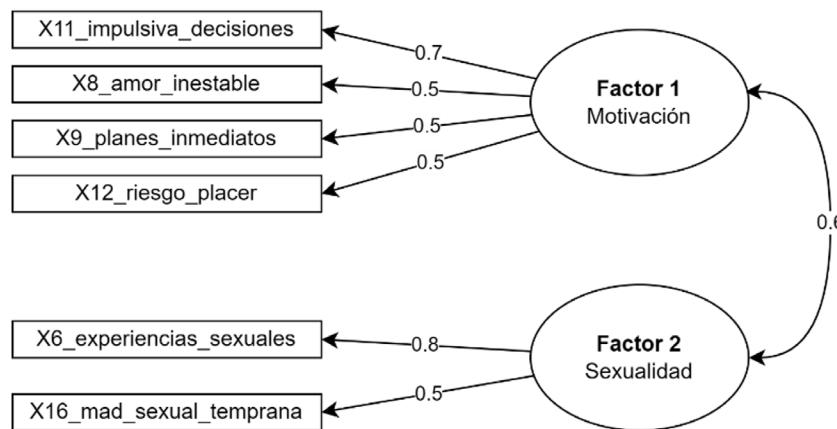


FIGURA 2. Diagrama del análisis factorial exploratorio con 2 factores.

Se emplearon 10 pliegues de validación para evitar el sobreajuste<sup>[43]</sup>, y las características consideradas como no informativas fueron eliminadas. De las 19 variables seleccionadas del instrumento K-SF-42, se eliminaron '*Familia se preocupa por bienestar*', '*Forma de ser familiares*' y '*Familia muestra afecto*'. Además, se agregaron dos variables referentes a la preocupación para complementar la información de las otras variables seleccionadas que son '*Bajo tensión*' y '*Preocupación excesiva*'. El resultado final después de utilizar RFE fueron 27 variables seleccionadas para el entrenamiento de los modelos.

### Balanceo de clases

En el conjunto de datos original existe un ligero desbalance en las tres clases a detectar de ansiedad, por lo que para mejorar el desempeño del modelo se decidió utilizar las técnicas de SMOTE y SMOTE-Tomek Links. SMOTE (*Synthetic Minority Over-sampling Technique*) es una técnica de sobremuestreo utilizada en problemas de desbalance de clases en aprendizaje automático donde se crean nuevas instancias sintéticas mediante interpolación entre el punto de datos original y sus vecinos cercanos<sup>[44]</sup>. Por otro lado, SMOTE Tomek Links es un enfoque híbrido que combina el método de SMOTE con los enlaces de Tomek (*Tomek Links*) para mejorar el balance de clases en un conjunto de datos, mientras se elimina ruido y ejemplos ambiguos<sup>[45]</sup>.

Los resultados del balanceo de clases con las técnicas de SMOTE y SMOTE-Tomek Links se muestran en la Figura 3. Utilizando la técnica de SMOTE, todas las clases se equilibran a 92 instancias, mientras que la técnica de SMOTE-Tomek Links modifica ligeramente las instancias de la clase de 'Control sano de ansiedad' donde incrementa a 83 y disminuye las instancias de 'Ansiedad no TAG' de 91 a 81, y de la clase 'TAG' de 74 a 72 instancias.

### Conjunto de datos

El conjunto de datos final está compuesto por 27 variables: (1) 2 variables sociodemográficas, (2) el factor K normalizado, (3) 6 ítems pertenecientes a los marcadores de estrategias de historia de vida, específicamente de los

marcadores de sexualidad y motivación, (4) 16 ítems del instrumento K-SF-42, y (5) dos ítems del instrumento PSWQ. Las variables están resumidas en la Tabla 4.

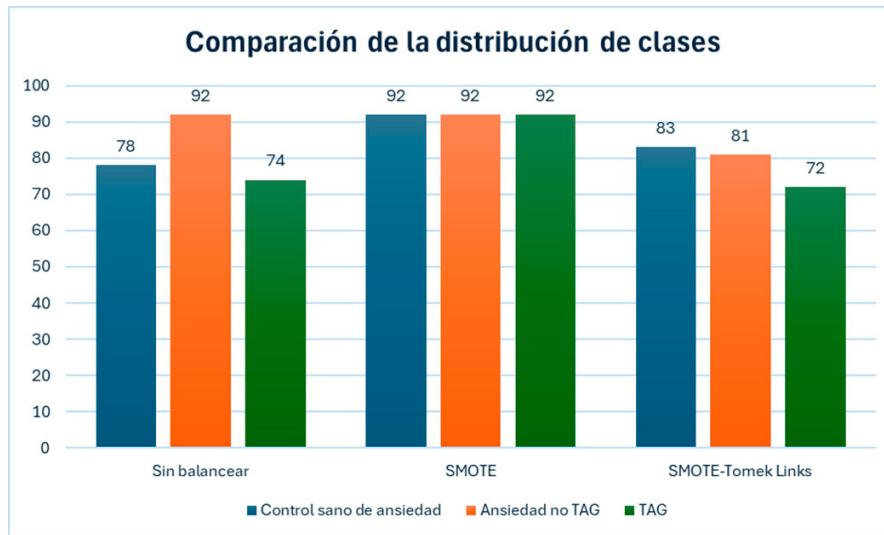


FIGURA 3. Distribución de las clases de instancias de acuerdo con la técnica de balanceo.

TABLA 4. Descripción general de las variables del conjunto de datos final.

Instrumento al que pertenece	Variable	Descripción
Sociodemográfico	Edad	Variable numérica que representa la edad del derechohabiente en años.
	Sexo	Variable binaria indicando el sexo del derechohabiente. '0' representa mujer, '1' representa hombre.
Marcadores de estrategias de historia de vida	Experiencias sexuales	Variable categórica con un rango de valores es de 1 a 5, donde 1 es 'No me describe en absoluto' y 5 es 'Me describe en absoluto'.
	Amor inestable	
	Planes inmediatos	
	Decisiones impulsivas	
	Riesgo placer	
	Madurez sexual temprana	
K-SF-42	Manera diferente	Variable categórica con un rango de valores de -3 a 3, donde -3 es 'Fuertemente en desacuerdo' y 3 es 'Fuertemente de acuerdo'.
	Positivo en peores situaciones	
	Bienestar parientes	
	Bienestar amigos	
	Contribuyo a causas	
	Pareja no interés	
	Marcar distancia	
	Fusionarme pareja	
	Demasiada intimidad	
	Pareja no intimidad	
	Atención mamá	Variable categórica con un rango de valores de 0 a 3, donde 0 es 'Nunca' y 3 es 'Mucho'.
	Atención papá	
	Enseñaba papá	
	Familia escuchar sentimientos	
	Ayuda amigos	Variable numérica obtenido como resultado del instrumento K-SF-42. El rango de valores es de 0 a 1.
	Hiciste bien amigos	
	Factor K normalizado	
PSWQ	Bajo tensión	Variable categórica con un rango de valores es de 1 a 5, donde 1 es 'No me describe en absoluto' y 5 es 'Me describe en absoluto'.
	Preocupación excesiva	
Variable objetivo	Categoría ansiedad	Variable categórica obtenida del puntaje total del cuestionario PSWQ. Las tres categorías posibles son 'Control sano de ansiedad', 'Ansiedad no TAG' y 'TAG'

### Desarrollo de un modelo de detección de trastorno de ansiedad generalizada

La comparación de diferentes modelos de aprendizaje automático y su desempeño es relevante para elegir el más adecuado para lograr el objetivo de la detección del trastorno. Se garantiza que el modelo seleccionado tenga tanto un rendimiento adecuado durante el entrenamiento y las pruebas, sino que también logre ser constante con ese rendimiento utilizando nuevos datos, lo que es crucial para obtener detecciones confiables<sup>[46]</sup>.

### Entrenamiento de modelos de aprendizaje automático

Los modelos de aprendizaje automático supervisados desarrollados en este trabajo son (1) Máquina de Soporte Vectorial (SVM), (2) *Random Forest*, (3) K Vecinos más Cercanos (KNN), (4) *Naïve Bayes*, (5) XGBoost, (6) AdaBoost, (7) Perceptrón Multicapa, y (8) Regresión Logística. Los algoritmos se entrenaron siguiendo la configuración inicial descrita en la Tabla 5. Para ello, se utilizó el 80% del total del conjunto de datos para el entrenamiento de los modelos, mientras que el 20% restante se destinó a la validación del modelo seleccionado.

Se realizó una comparación del rendimiento utilizando tres variantes del conjunto de datos: el conjunto original sin balancear, el conjunto balanceado mediante *SMOTE* y el conjunto balanceado con la técnica combinada *SMOTE-Tomek Links*. El entrenamiento de los modelos se llevó a cabo utilizando una técnica robusta de validación cruzada, la cual permite mejorar la capacidad de generalización del modelo frente a datos no vistos, al emplear diferentes particiones del conjunto de datos para entrenamiento y prueba en múltiples iteraciones<sup>[47]</sup>. En este estudio, se implementó validación cruzada estratificada con diez pliegues.

**TABLA 5. Configuración de hiperparámetros inicial de los algoritmos de aprendizaje supervisado.**

Modelo	Configuración de hiperparámetros
SVM	<code>kernel='rbf', C=10, gamma=0.01, class_weight='balanced'</code>
KNN	<code>n_neighbors=5, weights='uniform'</code>
Naïve Bayes	<code>var_smoothing=1e-09</code>
XGBoost	<code>eval_metric='mlogloss'</code>
Random Forest	<code>n_estimators=100, max_depth=None, max_features='sqrt'</code>
AdaBoost	<code>n_estimators=50, learning_rate=1.0</code>
Perceptrón Multicapa	<code>hidden_layer_sizes=(128, 64), max_iter=200, activation='relu', solver='adam', learning_rate_init=0.001</code>
Regresión Logística	<code>max_iter=500</code>

Después de realizar el entrenamiento de los modelos con la configuración inicial e identificar cuál conjunto de datos brindaba el mejor desempeño en los modelos, se realizó la optimización de hiperparámetros con la técnica de Búsqueda de Rejilla (*Grid Search*) para encontrar la mejor combinación de valores que brindaran una mejora en las métricas de desempeño. El espacio de búsqueda de hiperparámetros para cada algoritmo de aprendizaje se representa en la Tabla 6.

**TABLA 6. Espacio de búsqueda utilizado en Búsqueda de Rejilla para los modelos supervisados.**

Modelo	Espacio de búsqueda de hiperparámetros
SVM	<code>C: [0.1, 1, 10, 100], gamma: [0.001, 0.01, 0.1, 1], kernel: ['rbf', 'linear']</code>
KNN	<code>n_neighbors: [2, 3, 4, 5, 6], weights: ['uniform', 'distance'], p: [1, 2]</code>
Naïve Bayes	<code>var_smoothing: np.logspace(-11, -7, 5)</code>
XGBoost	<code>n_estimators: [100, 200, 300], max_depth: [3, 5, 7], learning_rate: [0.01, 0.1, 0.2]</code>
Random Forest	<code>n_estimators: [100, 200, 300], max_depth: [None, 10, 20], min_samples_split: [2, 5, 10]</code>
AdaBoost	<code>n_estimators: [50, 100, 200], learning_rate: [0.01, 0.1, 1.0]</code>
Perceptrón Multicapa	<code>hidden_layer_sizes: [(64,), (128,), (128, 64)], activation: ['relu', 'tanh'], alpha: [0.0001, 0.001, 0.01]</code>
Regresión Logística	<code>C: [0.01, 0.1, 1, 10]</code>

### Evaluación de modelos de aprendizaje automático

La evaluación de los modelos de aprendizaje es importante ya que nos permite comparar su rendimiento entre ellos y con los ya existentes a través de métricas que estiman si la predicción del modelo iguala o supera el desempeño de un profesional con experiencia<sup>[46]</sup>. Esta evaluación se realiza a través de métricas derivadas de una matriz de confusión. Esta matriz de confusión tabula la relación entre los resultados predichos y los resultados reales, obteniendo número de verdaderos positivos, falsos positivos, verdaderos negativos y falsos negativos<sup>[48]</sup>. En la Tabla 7 se presentan las métricas utilizadas para evaluar los modelos.

**TABLA 7. Resumen de las métricas de evaluación utilizadas.**

Métrica	Definición	Fórmula
<b>Exactitud</b>	Proporción de predicciones correctas entre todos los valores.	$\frac{VP + VN}{VP + VN + FP + FN}$
<b>Precisión</b>	Predicciones de verdaderos positivos divididas por todas las predicciones positivas.	$\frac{VP}{VP + FP}$
<b>Sensibilidad</b>	Predicciones de verdaderos positivos divididos por todos los valores positivos.	$\frac{VP}{VP + FN}$
<b>F1 score</b>	Media armónica de precisión y sensibilidad.	$\frac{2 \times \text{precisión} \times \text{sensibilidad}}{\text{precisión} + \text{sensibilidad}}$

*Verdaderos positivos (VP). Falsos positivos (FP). Falsos negativos (FN). Verdaderos negativos (VN).*

## RESULTADOS Y DISCUSIÓN

En esta sección se presentan los resultados obtenidos de la experimentación en este trabajo y se discuten los hallazgos en el contexto de la detección de trastornos de ansiedad implementando modelos de aprendizaje automático supervisado y los resultados de los análisis estadísticos.

Los resultados obtenidos del entrenamiento de los algoritmos seleccionados utilizando el conjunto de datos sin balancear se muestran en la Tabla 8. Los modelos de algoritmos de ensamble - *Random Forest* y *XGBoost* - son los que obtuvieron los mejores resultados en las métricas de exactitud, precisión, sensibilidad y F1 Score, obteniendo una exactitud del 78% y un F1 Score de 77%. El modelo KNN obtuvo un rendimiento del 53% en exactitud, demostrando un desempeño inferior al de los otros modelos desarrollados.

**TABLA 8. Resultados del entrenamiento con validación cruzada de los algoritmos supervisados utilizando el conjunto de datos no balanceado.**

Modelo	Exactitud	Precisión	Sensibilidad	F1 Score
SVM	0.67	0.69	0.67	0.67
KNN	0.53	0.57	0.53	0.52
Naive Bayes	0.74	0.76	0.74	0.73
<b>XGBoost</b>	<b>0.78</b>	<b>0.80</b>	<b>0.78</b>	<b>0.77</b>
<b>Random Forest</b>	<b>0.78</b>	<b>0.80</b>	<b>0.78</b>	<b>0.77</b>
AdaBoost	0.71	0.76	0.71	0.70
Perceptrón Multicapa	0.71	0.75	0.71	0.71
Regresión Logística	0.73	0.76	0.73	0.72

. Los resultados obtenidos del entrenamiento utilizando el conjunto de datos balanceado con SMOTE se muestran en la Tabla 9. Todos los modelos desarrollados presentaron mejoras en los valores de las métricas al ser balanceados mediante SMOTE con un incremento del rendimiento del 2 al 5%, donde específicamente XGBoost y Random Forest alcanzaron un F1-score del 80% o superior. En contraste, el rendimiento del algoritmo KNN presentó nuevamente un desempeño considerablemente bajo (inferior al 60%).

**TABLA 9. Resultados del entrenamiento con validación cruzada de los algoritmos supervisados utilizando el conjunto de datos balanceado con SMOTE.**

Modelo	Exactitud	Precisión	Sensibilidad	F1 Score
SVM	0.73	0.75	0.73	0.73
KNN	0.55	0.56	0.55	0.53
Naive Bayes	0.77	0.79	0.77	0.78
<b>XGBoost</b>	<b>0.83</b>	<b>0.84</b>	<b>0.83</b>	<b>0.82</b>
Random Forest	0.81	0.83	0.81	0.81
AdaBoost	0.76	0.81	0.76	0.76
Perceptrón Multicapa	0.76	0.77	0.76	0.76
Regresión Logística	0.77	0.79	0.79	0.78

Finalmente, los resultados obtenidos del entrenamiento de los algoritmos utilizando el conjunto de datos balanceado con SMOTE-Tomek Links se muestran en la Tabla 10. La mayoría de los algoritmos lograron un mejor desempeño en las métricas evaluadas en comparación con la técnica de SMOTE y el conjunto de datos sin balancear. Las métricas de desempeño presentaron valores iguales o superiores al 80%, destacando el modelo XGBoost que logró una exactitud y un F1 Score de 88%, seguido del modelo de Regresión Logística que obtuvo valores iguales o superiores al 84% en las métricas evaluadas. En el caso particular del modelo XGBoost, se tuvo una mejora del 10% o más en las métricas de desempeño al utilizar el balanceo SMOTE-Tomek Links.

**TABLA 10. Resultados del entrenamiento con validación cruzada de los algoritmos supervisados utilizando el conjunto de datos balanceado con SMOTE-Tomek Links.**

Modelo	Exactitud	Precisión	Sensibilidad	F1 Score
SVM	0.78	0.78	0.78	0.77
KNN	0.59	0.60	0.59	0.57
Naive Bayes	0.75	0.79	0.75	0.75
<b>XGBoost</b>	<b>0.88</b>	<b>0.89</b>	<b>0.88</b>	<b>0.88</b>
Random Forest	0.82	0.81	0.82	0.81
AdaBoost	0.82	0.85	0.82	0.82
Perceptrón Multicapa	0.79	0.80	0.79	0.78
Regresión Logística	0.85	0.86	0.85	0.84

Para optimizar los modelos que mostraron mejor desempeño en términos de exactitud, precisión, sensibilidad y F1 Score, los modelos se ajustaron mediante la técnica de optimización de hiperparámetros de Búsqueda de Rejilla (*Grid Search*) sobre los datos de entrenamiento. Los hiperparámetros óptimos obtenidos se muestran en la Tabla 11. Se evaluó el rendimiento de los modelos utilizando los hiperparámetros seleccionados por la técnica utilizando el conjunto de datos balanceado con SMOTE-Tomek Links para medir la mejora del desempeño (Tabla 12).

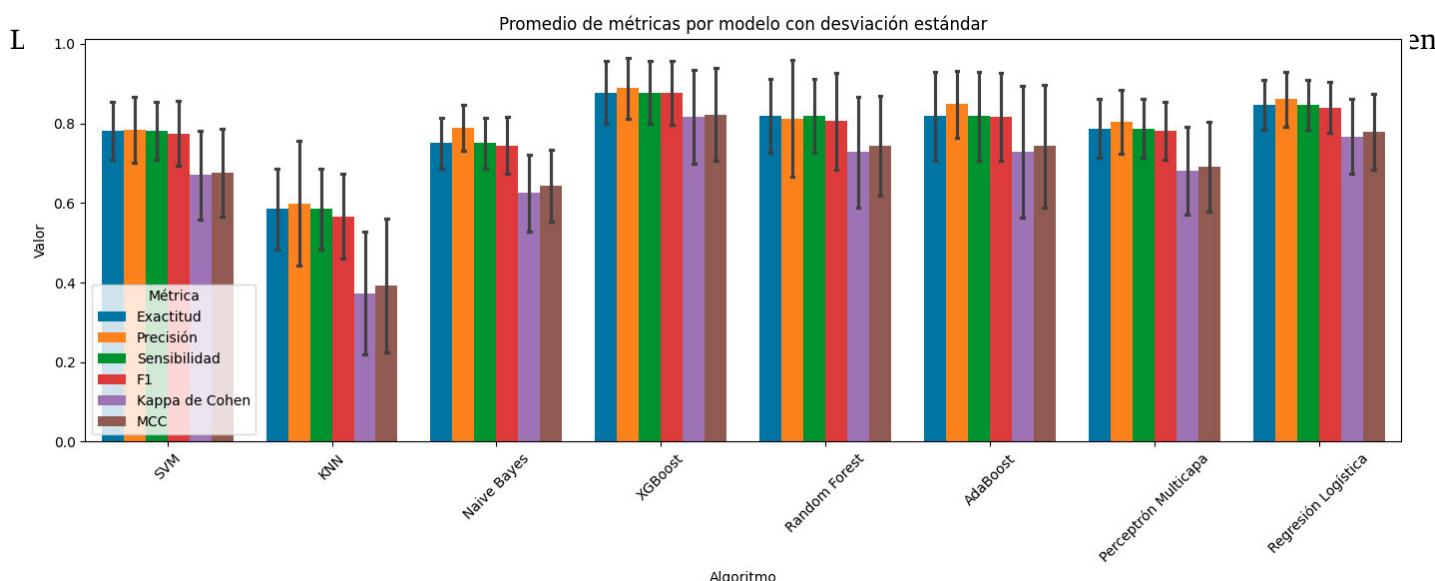
**TABLA 11. Configuración de hiperparámetros ideal obtenida con la técnica de Búsqueda de Rejilla (Grid Search).**

Modelo	Configuración de hiperparámetros
SVM	<code>kernel='linear', C=10, gamma=0.001, class_weight=None</code>
KNN	<code>n_neighbors=5, weights='distance', p=1</code>
Naive Bayes	<code>var_smoothing=1e-11</code>
XGBoost	<code>eval_metric="mlogloss", n_estimators= 200, max_depth= 5, learning_rate= 0.1</code>
Random Forest	<code>n_estimators=200, max_depth=None, max_features=None, min_samples_split: 2</code>
AdaBoost	<code>n_estimators=50, learning_rate=0.1</code>
Perceptrón Multicapa	<code>hidden_layer_sizes=(128, 64), max_iter=200, activation='tanh', solver='adam', learning_rate_init=0.001, alpha= 0.01</code>
Regresión Logística	<code>max_iter=500, C= 1</code>

**TABLA 12. Resultados del desempeño con validación cruzada de los algoritmos supervisados utilizando el conjunto de datos balanceado con SMOTE-Tomek Links v la configuración de hiperparámetros de Búsqueda de Rejilla (Grid Search).**

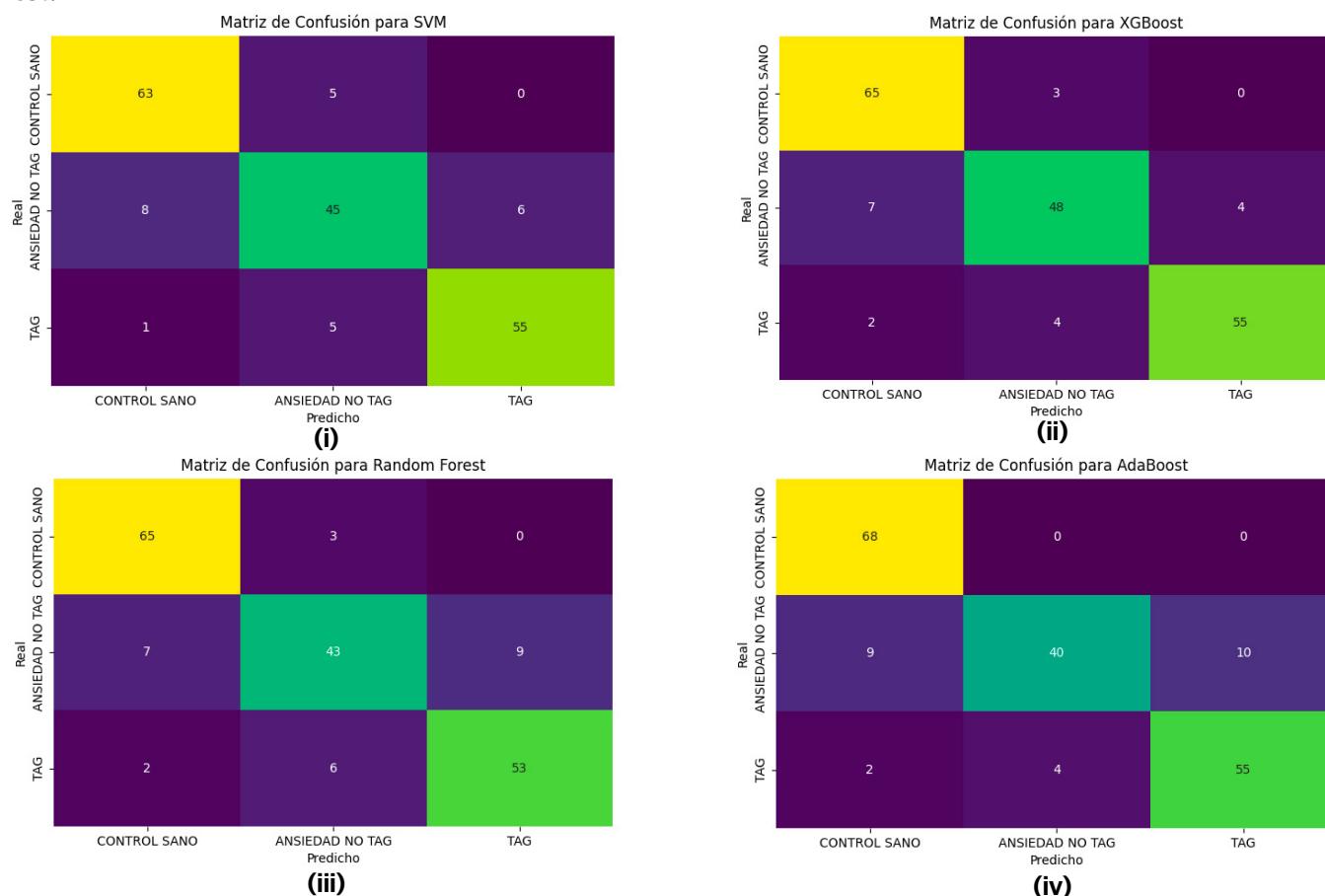
Modelo	Exactitud	Precisión	Sensibilidad	F1 Score
SVM	0.87	0.88	0.87	0.87
KNN	0.71	0.69	0.71	0.69
Naive Bayes	0.75	0.79	0.75	0.75
<b>XGBoost</b>	<b>0.89</b>	<b>0.91</b>	<b>0.89</b>	<b>0.89</b>
Random Forest	0.86	0.87	0.86	0.85
AdaBoost	0.87	0.88	0.87	0.86
Perceptrón Multicapa	0.83	0.84	0.83	0.83
Regresión Logística	0.85	0.86	0.85	0.84

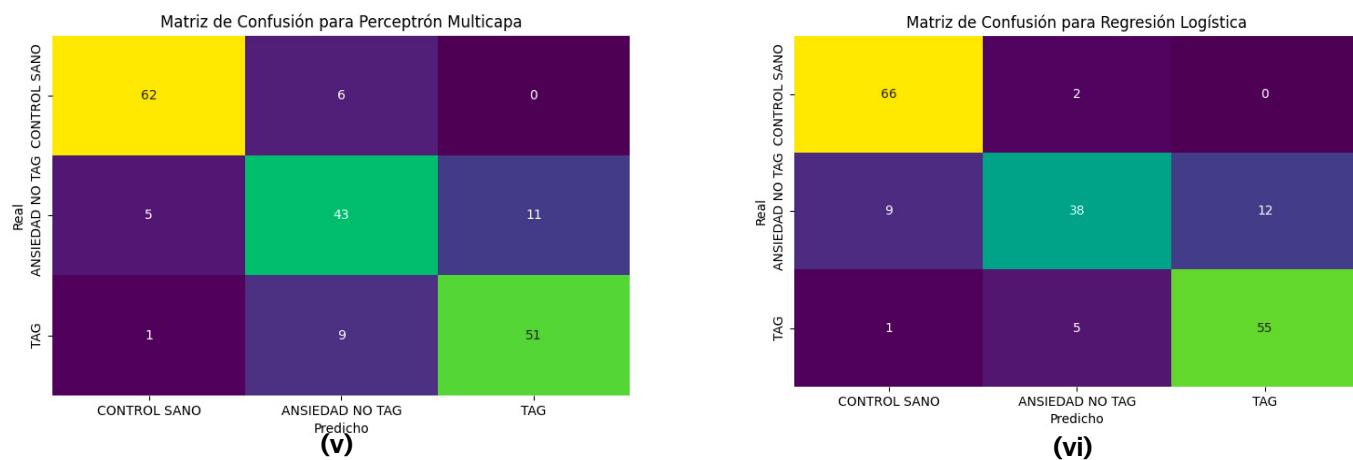
El modelo de XGBoost, balanceado mediante la técnica de SMOTE-Tomek Links y optimizado con Búsqueda de Rejilla, obtuvo el mejor desempeño en todas las métricas evaluadas: exactitud (89%), precisión (91%), sensibilidad (89%) y F1 Score (89%). En la Figura 4 se presenta una comparación del rendimiento de los diferentes modelos, incluyendo la variabilidad entre pruebas, donde se observa que XGBoost supera el 85 % en todas las métricas, con una baja desviación estándar, lo que lo posiciona como uno de los modelos más estables. En contraste, KNN y Naive Bayes mostraron un desempeño bajo e inestable, por lo que podrían descartarse para este caso. Otros algoritmos como SVM, Random Forest y Regresión Logística también alcanzaron resultados competitivos, con métricas iguales o superiores al 85 % y buena consistencia, por lo que representan alternativas viables para el problema abordado.

**FIGURA 4. Promedio de los resultados de la validación cruzada por métrica evaluada y la desviación estándar.**

Las matrices de confusión (Figura 5) de los modelos entrenados que obtuvieron resultados superiores al 80% en las cuatro métricas evaluadas permiten visualizar el desempeño de cada modelo en la clasificación de las tres clases: Control sano (68 instancias), Ansiedad no TAG (59 instancias), y TAG (61 instancias). El modelo XGBoost mostró un excelente rendimiento, con una alta tasa de aciertos en las tres clases. La clase “TAG” fue clasificada correctamente el 90% de los casos, mientras que para “Control sano” se tuvo un 95% de casos correctamente clasificados. La clase “Ansiedad no TAG” es la que tuvo un número mayor de casos incorrectamente clasificados ya que se confunden tanto con “Control sano” como con “TAG”. Al comparar todos los modelos, XGBoost, Regresión Logística, SVM y AdaBoost destacaron por su precisión en la clasificación de los casos de TAG.

Para comparar el desempeño de los modelos en cada una de las métricas evaluadas –exactitud, precisión, sensibilidad y F1 Score– se aplicó la prueba de *Friedman* sobre los resultados obtenidos en los 10 pliegues de validación cruzada, considerando los modelos entrenados con el conjunto de datos balanceado mediante *SMOTE-Tomek Links* y optimizados con búsqueda en rejilla. Se estableció un nivel de significancia de  $p < 0.05$ . Los resultados indicaron diferencias estadísticamente significativas en todas las métricas ( $p < 0.001$ ). Posteriormente, se realizaron pruebas post-hoc de comparación por pares utilizando la prueba de Nemenyi, con el objetivo de identificar en qué modelos radicaban las diferencias observadas. Los resultados mostraron que XGBoost obtuvo diferencias estadísticamente significativas en comparación con KNN y *Naive Bayes* en todas las métricas evaluadas, con  $p$ -valores inferiores a 0.05 (Tabla 13). En contraste, no se observaron diferencias significativas entre XGBoost y el resto de los modelos. Estos resultados sugieren que, aunque XGBoost se posiciona como uno de los modelos con mejor rendimiento general, su desempeño no difiere significativamente de otros algoritmos, como SVM y *Random Forest*.





**FIGURA 5.** Matrices de confusión de los modelos optimizados con Búsqueda de Rejilla que utilizan el conjunto de datos balanceado con SMOTE-Tomek Links. (i) Modelo SVM, (ii) Modelo XGBoost, (iii) Modelo Random Forest, (iv) Modelo AdaBoost, (v) Modelo Perceptrón Multicapa, (vi) Modelo Regresión Logística estandar.

**TABLA 13.** Resultados de las pruebas por pares con la prueba de Nemenyi del modelo XGBoost y los otros modelos implementados.

Modelos comparados	Exactitud (p-valor)	Precisión (p-valor)	Sensibilidad (p-valor)	F1 Score (p-valor)
XGBoost – KNN	0.014	0.026	0.014	0.010
XGBoost – Naive Bayes	0.012	0.035	0.012	0.026
XGBoost – SVM	0.998	1	0.998	1
XGBoost – Random Forest	0.907	0.936	0.907	0.958
XGBoost – AdaBoost	0.999	1	0.999	0.992
XGBoost – Regresión Logística	0.871	0.948	0.871	0.890
XGBoost – Perceptrón Multicapa	0.907	0.967	0.907	0.980

Nivel de significancia  $p < 0.05$

Al realizar la validación del modelo XGBoost se observó un buen desempeño (Tabla 14), alcanzando una exactitud de 85% y una exactitud balanceada de 83%, indicando una capacidad adecuada para clasificar correctamente las tres clases. Además, obtuvo un valor de precisión, sensibilidad y F1 Score del 85 %, lo que sugiere un equilibrio adecuado entre la identificación de los casos positivos y la precisión en sus predicciones positivas. La matriz de confusión (Figura 6) muestra que el modelo clasificó correctamente el 100% de los casos de “TAG”, mientras que la mayor confusión para el modelo es para la clase intermedia (Ansiedad no TAG) , lo cual es esperable dado que representa un punto intermedio en el continuo de severidad.

Además del buen desempeño observado en las métricas de validación, los resultados del modelo XGBoost propuesto en este estudio se destacan en comparación con investigaciones previas que han abordado la detección de ansiedad a partir de distintos tipos de datos y tamaños muestrales. En trabajos que emplean muestras menores a 1,000 instancias, los niveles de exactitud y sensibilidad suelen ser inferiores a los alcanzados en este estudio. Por ejemplo, en<sup>[27]</sup>, utilizando una muestra de 329 estudiantes y el inventario BAI, el mejor modelo (SVM) obtuvo una exactitud del 69.7 % y un AUC de 0.76. De manera similar, en<sup>[29]</sup>, con una muestra de 580 adolescentes, se alcanzó una exactitud balanceada del 62 %, con una sensibilidad del 53 % y un AUC de 0.69. Otros estudios como<sup>[33]</sup> y<sup>[34]</sup>, basados en muestras de 220 y 111 pacientes respectivamente, alcanzaron sensibilidades entre el 69 % y el 75 % y especificidades entre el 70 % y el 76 %. En contraste, el presente estudio reporta una exactitud del 85 % y una sensibilidad del 85 % en la clasificación multiclase de niveles de ansiedad.

TABLA 14. Resultados de la validación del modelo XGBoost desarrollado.

Modelo	Exactitud	Exactitud balanceada	Precisión	Sensibilidad	F1 Score
XGBoost	0.85	0.83	0.85	0.85	0.85

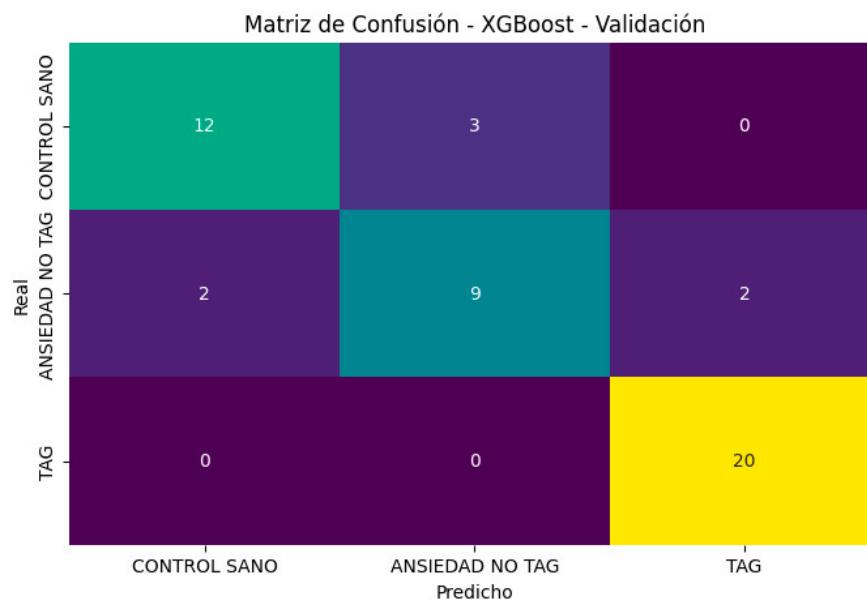


FIGURA 6. Matriz de confusión resultante de la validación del modelo XGBoost desarrollado.

La Figura 7 muestra la importancia de características según el criterio de frecuencia de uso (F score) en el modelo XGBoost. El F score representa cuántas veces se utilizó una característica para hacer una división en los árboles del modelo. La variable 'Preocupación excesiva' fue utilizada con mayor frecuencia para realizar divisiones en los árboles del modelo, destacándose como el principal predictor. Otras variables relevantes incluyeron 'Bajo tensión', 'Decisiones impulsivas' y 'Pareja no interés'. En conjunto, estos resultados sugieren que los aspectos relacionados con la preocupación, impulsividad y las relaciones interpersonales desempeñan un papel clave en la predicción realizada por el modelo.

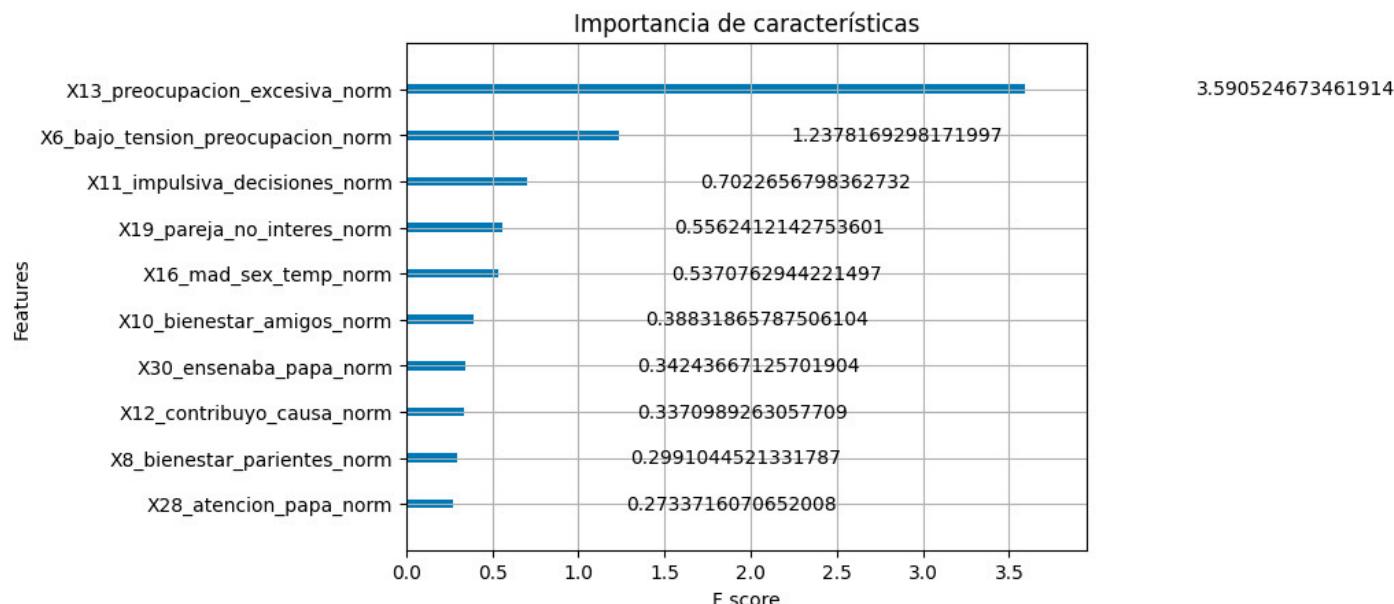


FIGURA 7. Importancia las 10 características más relevantes con el método 'gain' del modelo XGBoost.

Este hallazgo es consistente con lo reportado en la literatura. Se ha documentado que la preocupación, lejos de actuar como una estrategia de afrontamiento efectiva, puede funcionar como un mecanismo compensatorio que perpetúa las emociones negativas y reduce la capacidad de experimentar emociones positivas<sup>[49]</sup>. En el mismo sentido, la inseguridad en los vínculos afectivos ha sido asociada con mayores niveles de ansiedad y una menor capacidad para manejar el estrés<sup>[50]</sup>. Asimismo, la maduración precoz se ha vinculado con un mayor riesgo de depresión, ansiedad e inicio temprano de la actividad sexual<sup>[6]</sup>, lo cual refuerza la importancia de variables como 'Madurez sexual temprana' en las predicciones del modelo. Finalmente, aunque la impulsividad no suele ser una variable principal en el estudio de TAG, investigaciones recientes señalan que las personas con TAG presentan niveles más altos de impulsividad, y que esta podría estar mediada más por el afecto negativo que por la preocupación en sí misma<sup>[51][52]</sup>.

## CONCLUSIONES

En los últimos años, el área de la salud mental ha experimentado avances significativos en la aplicación de modelos de aprendizaje automático para la detección de trastornos mentales, incluyendo la ansiedad. Los resultados de este estudio confirman la eficacia de los modelos supervisados en esta tarea, destacando particularmente a XGBoost por su rendimiento general. No obstante, su desempeño no difirió significativamente del alcanzado por otros algoritmos robustos como *Random Forest* y SVM, lo que sugiere que diferentes enfoques pueden ser válidos siempre que se apliquen técnicas adecuadas de optimización y preparación de los datos. La combinación de técnicas de balanceo como *SMOTE-Tomek Links*, la optimización de hiperparámetros mediante búsqueda en rejilla y el uso de algoritmos de ensamble contribuyó significativamente a mejorar la detección de distintos niveles de ansiedad en comparación con enfoques tradicionales. En particular, el modelo XGBoost alcanzó una exactitud del 89% en el entrenamiento y del 85% en la validación para la detección de TAG, siendo el modelo que demostró mayor estabilidad y consistencia con este conjunto de datos.

Un aporte de este trabajo fue la incorporación de variables relacionadas con las estrategias de historia de vida, lo que permitió desarrollar un modelo más robusto y explicativo. Este enfoque mostró la importancia de factores emocionales y relacionales en la manifestación del TAG. Las variables como la preocupación excesiva, la impulsividad, la madurez sexual temprana y la calidad de los vínculos afectivos se identificaron como elementos clave en la predicción, en concordancia con lo reportado en la literatura científica.

El modelo validado mantuvo un desempeño consistente en las etapas de entrenamiento y validación, lo que indicaría una adecuada capacidad de generalización que necesita ser comprobada con estudios subsecuentes. En general, los hallazgos de este estudio evidencian que la integración de modelos de aprendizaje automático con variables de historia de vida representa una estrategia prometedora para el diagnóstico automatizado de trastornos mentales. Se recomienda para investigaciones futuras ampliar el número y diversidad de las variables analizadas, así como fortalecer la representatividad de las muestras, con el fin de incrementar la precisión y aplicabilidad clínica de estos modelos en contextos de salud mental, donde sea posible identificar personas con probabilidad de desarrollar TAG desde una temprana edad.

### Declaración ética

El estudio fue evaluado y aprobado por el Comité de Ética en Investigación y el Comité Local de Investigación del

IMSS de Morelos, obteniendo el número de registro R-2024-1701-029.

### Conflictos de interés

Los autores declaran que no tienen conflictos de intereses.

### REFERENCIAS

- [1] “Clasificación Internacional de Enfermedades, undécima revisión (CIE-11), Organización Mundial de la Salud (OMS)”. enero de 2024. [En línea]. Disponible en: <https://icd.who.int/browse/2024-01/mms/es#1712535455>.
- [2] P. Gaitán-Rossi, V. Pérez-Hernández, M. Vilar-Compte, y G. Teruel-Belismelis, “Prevalencia mensual de trastorno de ansiedad generalizada durante la pandemia por Covid-19 en México”, Salud Pública Mex, vol. 63, núm. 4, pp. 478-485, jun. 2021, doi: <https://doi.org/10.21149/12257>.
- [3] S. Graham et al., “Artificial Intelligence for Mental Health and Mental Illnesses: an Overview”, Curr Psychiatry Rep, vol. 21, núm. 11, p. 116, nov. 2019, doi: <https://doi.org/10.1007/s11920-019-1094-0>.
- [4] M. Del Giudice y B. J. Ellis, Evolutionary foundations of developmental psychopathology., 3a ed. John Wiley & Sons, 2016.
- [5] P. S. Kavanagh y B. L. Kahl, “Life History Theory”, en Encyclopedia of Evolutionary Psychological Science, V. Weekes-Shackelford, T. K. Shackelford, y V. A. Weekes-Shackelford, Eds., Cham: Springer International Publishing, 2016, pp. 1-12. doi: [https://doi.org/10.1007/978-3-319-16999-6\\_1914-1](https://doi.org/10.1007/978-3-319-16999-6_1914-1).
- [6] B. L. Kahl, P. S. Kavanagh, y D. H. Gleaves, “Testing a life history model of psychopathology: A replication and extension”, Curr Psychol, vol. 41, núm. 9, pp. 6233-6246, sep. 2022, doi: <https://doi.org/10.1007/s12144-020-01062-y>.
- [7] K. J. Chua, A. W. Lukaszewski, D. M. Grant, y O. Sng, “Human Life History Strategies: Calibrated to External or Internal Cues?”, Evol Psychol, vol. 15, núm. 1, p. 1474704916677342, ene. 2017, doi: <https://doi.org/10.1177/1474704916677342>.
- [8] H. J. Lu, A. T. Yang, Y. Y. Liu, N. Zhu, y L. Chang, “Being Cared for and Growing Up Slowly: Parenting Slows Human Life History”, Parenting, vol. 23, núm. 2, pp. 140-158, abr. 2023, doi: <https://doi.org/10.1080/15295192.2023.2243500>.
- [9] A. Burkov, The Hundred-Page Machine Learning Book, 1a ed. 2019.
- [10] G. Cho, J. Yim, Y. Choi, J. Ko, y S.-H. Lee, “Review of Machine Learning Algorithms for Diagnosing Mental Illness”, Psychiatry Investig, vol. 16, núm. 4, pp. 262-269, abr. 2019, doi: <https://doi.org/10.30773/pi.2018.12.21.2>.
- [11] J. Chung y J. Teo, “Single classifier vs. ensemble machine learning approaches for mental health prediction”, Brain Inform, vol. 10, núm. 1, p. 1, ene. 2023, doi: <https://doi.org/10.1186/s40708-022-00180-6>.
- [12] S. Shalev-Shwartz y S. Ben-David, Understanding machine learning: from theory to algorithms. New York: Cambridge university press, 2014.
- [13] C. Ordóñez, “Comparing association rules and decision trees for disease prediction”, en Proceedings of the international workshop on Healthcare information and knowledge management, Arlington Virginia USA: ACM, nov. 2006, pp. 17-24. doi: <https://doi.org/10.1145/1183568.1183573>.
- [14] T. Chen y C. Guestrin, “XGBoost: A Scalable Tree Boosting System”, en Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco California USA: ACM, ago. 2016, pp. 785-794. doi: <https://doi.org/10.1145/2939672.2939785>.
- [15] T. Hastie, R. Tibshirani, y J. Friedman, The Elements of Statistical Learning. en Springer Series in Statistics. New York, NY: Springer New York, 2009. doi: <https://doi.org/10.1007/978-0-387-84858-7>.
- [16] Y. Freund y R. E. Schapire, “A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting”, Journal of Computer and System Sciences, vol. 55, núm. 1, pp. 119-139, ago. 1997, doi: <https://doi.org/10.1006/jcss.1997.1504>.
- [17] I. Goodfellow, Y. Bengio, y A. Courville, Deep Learning. MIT Press, 2016.
- [18] D. E. Rumelhart, G. E. Hinton, y R. J. Williams, “Learning representations by back-propagating errors”, Nature, vol. 323, núm. 6088, pp. 533-536, oct. 1986, doi: <https://doi.org/10.1038/323533a0>.
- [19] R. A. Fisher, “Statistical Methods for Research Workers”, en Breakthroughs in Statistics, S. Kotz y N. L. Johnson, Eds., en Springer Series in Statistics., New York, NY: Springer New York, 1992, pp. 66-70. doi: [https://doi.org/10.1007/978-1-4612-4380-9\\_6](https://doi.org/10.1007/978-1-4612-4380-9_6).
- [20] W. H. Kruskal y W. A. Wallis, “Use of Ranks in One-Criterion Variance Analysis”, Journal of the American Statistical Association, vol. 47, núm. 260, pp. 583-621, dic. 1952, doi: <https://doi.org/10.1080/01621459.1952.10483441>.
- [21] M. Friedman, “The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance”, Journal of the American Statistical Association, vol. 32, núm. 200, pp. 675-701, dic. 1937, doi: <https://doi.org/10.1080/01621459.1937.10503522>.
- [22] Nemenyi, P, “Distribution-free Multiple Comparisons”, Ph.D. dissertation, Princeton University, 1963.
- [23] H. Alharthi, “Predicting the level of generalized anxiety disorder of the coronavirus pandemic among college age students using artificial intelligence technology”, en 2020 19th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES), Xuzhou, China: IEEE, oct. 2020, pp. 218-221. doi: <https://doi.org/10.1109/DCABES50732.2020.00064>.
- [24] M. Wang, L. L. Richmond, J. L. Schleider, B. D. Nelson, y C. C. Luhmann, “Predicting internalizing symptoms with machine learning: identifying individuals that need care”, Journal of American College Health, pp. 1-10, nov. 2023, doi: <https://doi.org/10.1080/07448481.2023.2277185>.

- [25] Md. I. H. Nayan et al., “Comparison of the Performance of Machine Learning-based Algorithms for Predicting Depression and Anxiety among University Students in Bangladesh: A Result of the First Wave of the COVID-19 Pandemic”, *Asian Journal of Social Health and Behavior*, vol. 5, núm. 2, pp. 75-84, abr. 2022, doi: [https://doi.org/10.4103/shb.shb\\_38\\_22](https://doi.org/10.4103/shb.shb_38_22).
- [26] J. Maitre, C. Bergeron-Leclerc, D. Maltais, y S. Gaboury, “Investigating anxiety levels in the Quebec university community during the COVID-19 pandemic using machine learning and data exploration techniques”, *Multimed Tools Appl*, vol. 82, núm. 29, pp. 46109-46127, dic. 2023, doi: <https://doi.org/10.1007/s11042-023-16096-0>.
- [27] M. Mahalingam et al., “A Machine Learning Study to Predict Anxiety on Campuses in Lebanon”, en *Studies in Health Technology and Informatics*, J. Mantas, P. Gallos, E. Zoulias, A. Hasman, M. S. Househ, M. Charalampidou, y A. Magdalinos, Eds., IOS Press, 2023. doi: <https://doi.org/10.3233/SHTI230430>.
- [28] A. Daza, Arroyo-Paz, J. Bobadilla, O. Apaza, y J. Pinto, “Stacking ensemble learning model for predict anxiety level in university students using balancing methods”, *Informatics in Medicine Unlocked*, vol. 42, p. 101340, 2023, doi: <https://doi.org/10.1016/j.imu.2023.101340>.
- [29] A. V. Chavanne et al., “Anxiety onset in adolescents: a machine-learning prediction”, *Mol Psychiatry*, vol. 28, núm. 2, pp. 639-646, feb. 2023, doi: <https://doi.org/10.1038/s41380-022-01840-z>.
- [30] M. D. Nemesure, M. V. Heinz, R. Huang, y N. C. Jacobson, “Predictive modeling of depression and anxiety using electronic health records and a novel machine learning approach with artificial intelligence”, *Sci Rep*, vol. 11, núm. 1, p. 1980, ene. 2021, doi: <https://doi.org/10.1038/s41598-021-81368-4>.
- [31] R. Pal et al., “A Novel Smart Belt for Anxiety Detection, Classification, and Reduction Using IIoMT on Students’ Cardiac Signal and MSY”, *Bioengineering*, vol. 9, núm. 12, p. 793, dic. 2022, doi: <https://doi.org/10.3390/bioengineering9120793>.
- [32] J. Wang et al., “Analysis of Risk Factors and Diagnosis for Anxiety Disorder in Older People with the Aid of Artificial Intelligence: Observational Study”, en 2023 31st Irish Conference on Artificial Intelligence and Cognitive Science (AICS), Letterkenny, Ireland: IEEE, dic. 2023, pp. 1-8. doi: <https://doi.org/10.1109/AICS60730.2023.10470782>.
- [33] É. D. M. Silveira et al., “Decoding rumination: A machine learning approach to a transdiagnostic sample of outpatients with anxiety, mood and psychotic disorders”, *Journal of Psychiatric Research*, vol. 121, pp. 207-213, feb. 2020, doi: <https://doi.org/10.1016/j.jpsychires.2019.12.005>.
- [34] T. Richter, B. Fishbain, E. Fruchter, G. Richter-Levin, y H. Okon-Singer, “Machine learning-based diagnosis support system for differentiating between clinical anxiety and depression disorders”, *Journal of Psychiatric Research*, vol. 141, pp. 199-205, sep. 2021, doi: <https://doi.org/10.1016/j.jpsychires.2021.06.044>.
- [35] A. J. Figueiredo et al., “The K-SF-42: A New Short Form of the Arizona Life History Battery”, *Evol Psychol*, vol. 15, núm. 1, p. 147470491667627, ene. 2017, doi: <https://doi.org/10.1177/1474704916676276>.
- [36] M. Del Giudice, “The Life History Framework and the FSD Model”, en *Evolutionary Psychopathology*, M. Del Giudice, Ed., Oxford University Press, 2018, pp. 153-192. doi: <https://doi.org/10.1093/med-psych/9780190246846.003.0006>.
- [37] T. J. Meyer, M. L. Miller, R. L. Metzger, y T. D. Borkovec, “Development and validation of the penn state worry questionnaire”, *Behaviour Research and Therapy*, vol. 28, núm. 6, pp. 487-495, 1990, doi: [https://doi.org/10.1016/0005-7967\(90\)90135-6](https://doi.org/10.1016/0005-7967(90)90135-6).
- [38] M. Grebovic, L. Filipovic, I. Katnic, M. Vukotic, y T. Popovic, “Machine Learning Models for Statistical Analysis”, *IAJIT*, vol. 20, núm. 3A, 2023, doi: <https://doi.org/10.34028/ijait/20/3A/8>.
- [39] X. Cheng, “A Comprehensive Study of Feature Selection Techniques in Machine Learning Models”, *Ins. Comput. Signal Syst.*, vol. 1, núm. 1, pp. 65-78, nov. 2024, doi: <10.70088/xpf2b276>.
- [40] P. Rawat, “Importance of Feature Selection in Model Accuracy”, 2021. [En línea]. Disponible en: <https://api.semanticscholar.org/CorpusID:236907306>.
- [41] C. D. Stapleton, “Basic Concepts and Procedures of Confirmatory Factor Analysis.”, 1997. [En línea]. Disponible en: <https://api.semanticscholar.org/CorpusID:58944126>.
- [42] P. Misra y A. S. Yadav, “Improving the Classification Accuracy using Recursive Feature Elimination with Cross-Validation”, vol. 11, núm. 3, pp. 659-665, may 2020.
- [43] L. A. Yates, Z. Aandahl, S. A. Richards, y B. W. Brook, “Cross validation for model selection: A review with examples from ecology”, *Ecological Monographs*, vol. 93, núm. 1, p. e1557, feb. 2023, doi: <https://doi.org/10.1002/ecm.1557>.
- [44] N. V. Chawla, K. W. Bowyer, L. O. Hall, y W. P. Kegelmeyer, “SMOTE: Synthetic Minority Over-sampling Technique”, *jair*, vol. 16, pp. 321-357, jun. 2002, doi: <https://doi.org/10.1613/jair.953>.
- [45] M. Zeng, B. Zou, F. Wei, X. Liu, y L. Wang, “Effective prediction of three common diseases by combining SMOTE with Tomek links technique for imbalanced medical data”, en 2016 IEEE International Conference of Online Analysis and Computing Science (ICOACS), Chongqing, China: IEEE, may 2016, pp. 225-228. doi: <https://doi.org/10.1109/ICOACS.2016.7563084>.
- [46] O. Rainio, J. Teuho, y R. Klén, “Evaluation metrics and statistical tests for machine learning”, *Sci Rep*, vol. 14, núm. 1, p. 6086, mar. 2024, doi: <https://doi.org/10.1038/s41598-024-56706-x>.
- [47] X. Zeng y T. R. Martinez, “Distribution-balanced stratified cross-validation for accuracy estimation”, *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 12, núm. 1, pp. 1-12, ene. 2000, doi: <https://doi.org/10.1080/095281300146272>.
- [48] H. H. Rashidi, S. Albahra, S. Robertson, N. K. Tran, y B. Hu, “Common statistical concepts in the supervised Machine Learning arena”, *Front. Oncol.*, vol. 13, p. 1130229, feb. 2023, doi: <https://doi.org/10.3389/fonc.2023.1130229>.

- [49] C. Gale y O. Davidson, “Generalised anxiety disorder”, BMJ, vol. 334, núm. 7593, pp. 579-581, mar. 2007, doi: <https://doi.org/10.1136/bmj.39133.559282.be>.
- [50] Z. Dodd, S. Driver, A. Warren, S. Riggs, y M. Clark, “Effects of Adult Romantic Attachment and Social Support on Resilience and Depression in Individuals with Spinal Cord Injuries”, Topics in Spinal Cord Injury Rehabilitation, vol. 21, núm. 2, pp. 156-165, mar. 2015, doi: <https://doi.org/10.1310/sci2102-156>.
- [51] A. Del Carlo et al., “Different measures of impulsivity in patients with anxiety disorders: A case control study”, Psychiatry Research, vol. 197, núm. 3, pp. 231-236, may 2012, doi: <https://doi.org/10.1016/j.psychres.2011.09.020>.
- [52] R. Ferreira-Garcia, C. Gitahy Falcão Faria, A. E. Nardi, y R. C. da Rocha Freire, “Negative Affect Mediates Impulsivity in Generalized Anxiety Disorder”, Psychopathology, vol. 52, núm. 6, pp. 327-333, 2019, doi: <https://doi.org/10.1159/000503395>.