

Análisis de depresión en adolescentes de Básica Secundaria y Media usando algoritmos de *Machine Learning*

Analysis of depression in High School and Middle Basic adolescents using Machine Learning algorithms



ISSN 0124-0137
EISSN 2027-212X

ARTÍCULO DE
INVESTIGACIÓN
Copyright © 2024
by PsicoGente

Correspondencia de autores:

ymsaavedrag@ut.edu.co
jsalinas@lamolina.edu.pe
mari_apilar_garcia@hotmail.com

Recibido: 23-11-23
Aceptado: 28-08-24
Publicado: 01-09-24

Yuri Marcela García Saavedra 
Universidad del Tolima - Tolima, Colombia.

Jesús Walter Salinas Flores 
Universidad Nacional Agraria La Molina - Lima, Perú.

María Del Pilar García Saavedra 
Institución Educativa Técnica Alfonso
Palacio Rudas - Tolima, Colombia

Resumen:

Objetivo: Identificar los factores determinantes relacionados con la ansiedad, el estrés y la depresión en adolescentes de sexto a undécimo grado de una institución educativa del departamento del Tolima, Colombia.

Método: Se aplicó el Inventario Sistemático Cognoscitivista para el estudio del estrés académico (Inventario SISCO) a una muestra de 553 estudiantes cuyo rango de edad estuvo entre 11 y 20 años (media=17,12 años, DE=0,766). Se emplearon diferentes algoritmos de Machine Learning para clasificación debido a que la variable target es de respuesta binaria, complementando con el uso de métodos de ensamble.

Resultados: Al preguntarle a los estudiantes si alguna vez habían tenido sentimientos persistentes de tristeza, sensación de vacío o pérdida de interés en las actividades cotidianas (indicativas de depresión), el 76,17 % manifestaron que sí mientras que el 23,83 % respondió que no. La probabilidad de detectar correctamente a los estudiantes con depresión fue de 0,845, mientras que para los estudiantes que no manifiestan tener depresión fue de 0,863. Las principales variables que mejor discriminaron entre estudiantes que afirmaron tener síntomas de depresión y aquellos que no, fueron: fatiga crónica, dificultad para tomar decisiones, ansiedad, falta de apetito, conflictos interpersonales, actividades evaluativas y somnolencia.

Conclusiones: Se recalca la importancia de la detección temprana, la utilización de herramientas de Machine Learning, y la necesidad de enfoques preventivos y personalizados para abordar la ansiedad, el estrés y la depresión en adolescentes.

Palabras clave: Adolescentes, depresión, algoritmos de clasificación, Machine Learning, métodos de ensamble.

Abstract

Objective: To identify the determining factors associated with anxiety, stress, and depression in adolescents from sixth to eleventh grade at an educational institution in the department of Tolima, Colombia.

Method: The Systemic Cognitive Inventory for the Study of Academic Stress (SISCO Inventory) was applied to a sample of 553 students aged 11 to 20 (mean = 17,12 years, SD = 0,766). Various machine learning classification algorithms were employed since the target variable was binary, complemented by ensemble methods.

Results: When students were asked whether they had ever experienced persistent feelings of sadness, emptiness, or loss of interest in daily activities (indicative of depression), 76,17 % reported "yes," while 23,83 % answered "no." The probability of correctly detecting students with depression was 0,845, while for students without depression, it was 0,863. The main variables that best discriminated between students reporting depressive symptoms and those who did not include chronic fatigue, difficulty making decisions, anxiety, lack of appetite, interpersonal conflicts, evaluative activities, and drowsiness.

Cómo citar este artículo (APA):

García Saavedra, Y. M., Salinas Flores, J. W., & García Saavedra, M. del P. 2024. Análisis de depresión en adolescentes de Básica Secundaria y Media usando algoritmos de Machine Learning. *PsicoGente* 27(52), 1-27. <https://doi.org/10.17081/psico.27.52.6981>

Conclusions: The findings underscore the importance of early detection, the application of machine learning tools, and the need for preventive and personalized approaches to address anxiety, stress, and depression in adolescents.

Keywords: Adolescents, depression, classification algorithms, machine learning, ensemble methods.

1. INTRODUCCIÓN

Algunos autores se refieren a la ansiedad y la depresión como emociones o estados de ánimo que, según su intensidad o frecuencia, pueden constituir síntomas psicopatológicos (Erazo & Martínez, 2024; Martínez de Velasco *et al.*, 2023; Leonangeli *et al.*, 2024; Instituto Nacional de la Salud Mental [INSM], 2024). La ansiedad se considera un estado emocional desagradable, que se manifiesta a nivel cognitivo, fisiológico y motor (Rodríguez *et al.*, 2022; Hervías *et al.*, 2022; Neff, 2022). Santomauro *et al.* (2021) indicaron que la pandemia de COVID-19 en 2020 provocó un considerable incremento en los casos de ansiedad y trastornos depresivos a nivel mundial. Se estima que el 3,8 % de la población sufre de depresión, incluyendo un 5 % en adultos (4 % en hombres y 6 % en mujeres) y un 5,7 % en personas mayores de 60 años. En total, aproximadamente 280 millones de personas en el mundo padecen depresión (Organización Mundial de la Salud [OMS], 2023). Además, cerca de uno de cada siete adolescentes entre 10 y 19 años, es decir, el 14 %, presenta algún tipo de trastorno mental. Estas condiciones representan el 15 % de la carga global de morbilidad en adolescentes, siendo la depresión, la ansiedad y los trastornos de comportamiento las principales causas de enfermedad y discapacidad en esta etapa. Los trastornos de ansiedad, que pueden manifestarse como ataques de pánico o preocupación excesiva, son especialmente comunes y suelen afectar más a los adolescentes mayores. Se estima que un 4,4 % de los adolescentes entre 10 y 14 años y un 5,5 % de los de 15 a 19 años experimentan ansiedad, mientras que el 1,4 % de los adolescentes de 10 a 14 años y el 3,2 % de los de 15 a 19 años sufren de depresión (Organización Mundial de la Salud [OMS], 2024).

Para los adultos que presentan trastornos afectivos, de ansiedad, y relacionados con el consumo de sustancias, tanto graves como moderados, se estima que la brecha terapéutica mediana es del 47,2 % en América del Norte y del 77,9 % en América Latina y el Caribe. En cuanto a la brecha para la depresión es de 73,9 % y para el alcohol es de 85,1 % (Organización Panamericana de la Salud [OPS], 2018; Meylan *et al.*, 2024). En Colombia, la depresión representa la segunda mayor causa de carga de enfermedad, cerca del 44,7 % de niños y niñas muestran indicios de algún trastorno de salud mental. En la adolescencia, entre los 12 y 18 años, los problemas más comunes incluyen

la ansiedad, la fobia social y la depresión. Además, la ideación suicida afecta al 6,6 % de esta población, siendo más frecuente en mujeres (7,4 %) que en hombres (5,7 %). En la adultez, se estima que alrededor del 6,7 % de las personas ha experimentado algún tipo de trastorno afectivo (Ministerio de Salud y Protección Social, 2022).

Considerando la evidencia del incremento de los trastornos de ansiedad y depresión en la población colombiana, es fundamental reconocer la salud mental como un aspecto clave en el desarrollo adecuado de los adolescentes. Es esencial ofrecer herramientas efectivas para el manejo y control de estos trastornos, con el objetivo de prevenir efectos más graves en esta etapa de la vida.

1.1. Factores de Riesgo que Impactan la Salud Mental de los Adolescentes

Existen múltiples factores que afectan la salud mental. A medida que aumentan los factores de riesgo en la vida de los adolescentes, también se incrementa el impacto potencial en su bienestar mental. Entre estos factores se encuentran el estrés característico de esta etapa de la vida, la presión ejercida por el entorno social de los compañeros, así como el proceso de búsqueda de su identidad personal (Erazo & Martínez, 2024; Güemes-Hidalgo *et al.*, 2017; Pardo *et al.*, 2004), la influencia de los medios de comunicación y la imposición de normas de género como las relacionadas con la pureza o modestia para las mujeres o los mensajes que promueven actitudes más agresivas o dominantes, para los adolescentes varones, incluso, se espera que las mujeres elijan profesiones consideradas *femeninas*, mientras que a los hombres se les sugiere que sigan carreras de mayor prestigio o de carácter técnico. Todos estos aspectos pueden agravar la discrepancia entre la realidad que vive el adolescente y sus percepciones o aspiraciones de cara al futuro (Martín *et al.*, 2022; Ortiz-Hernández & Valencia-Valero, 2015). La salud mental de los adolescentes también está determinada por factores como la calidad de su entorno familiar y sus interacciones con los compañeros (Alcindor-Huelva *et al.*, 2019). Además, la violencia, una crianza excesivamente rígida y las dificultades socioeconómicas pueden tener un impacto negativo significativo (Bustos & Russo, 2018; Martín *et al.*, 2022; Salazar *et al.*, 2016).

Estos factores de riesgo podrían llevar a los jóvenes a experimentar trastornos depresivos y/o de ansiedad, sintiéndose frustrados por no contar con mecanismos de defensa adecuados para hacerles frente (Organización

Mundial de la Salud [OMS], 2023). Algunos adolescentes, entre los 12 y 17 años de edad, corren un mayor riesgo de sufrir trastornos de salud mental, como dependencia al alcohol o las drogas, problemas derivados del acoso escolar, adicción al internet, trastorno de personalidad antisocial, esquizofrenia, hiperactividad e impulsividad (Erazo & Martínez, 2024; Leonangeli *et al.*, 2024; Lopez-Vega *et al.*, 2020; Ministerio de Salud y Protección Social, 2015; Ursul *et al.*, 2022). Estas condiciones pueden ser resultado de sus circunstancias de vida, situaciones de estigmatización, discriminación, exclusión, o la falta de acceso a servicios y apoyo de calidad (Bustos & Russo, 2018; Ministerio de Salud y Protección Social, 2022; Ortiz-Hernández & Valencia-Valero, 2015). Esto incluye a los jóvenes con discapacidades, embarazos tempranos en adolescentes, enfermedades crónicas, así como a aquellos que no cuentan con una vivienda estable o carecen de recursos económicos para subsistir (Bonilla, 2009; Martín *et al.*, 2022; Mora & Hernández, 2015; Nickson *et al.* 2023; Pardo *et al.*, 2004; Rodríguez *et al.* 2012; Ursul *et al.*, 2022).

1.2. Factores Estresantes en el entorno académico y sus efectos en los estudiantes

En un entorno académico, se pueden identificar diversos factores que generan estrés en los estudiantes. Entre estos factores se encuentran la competencia con sus pares, la carga excesiva de trabajos, las características personales del docente, las evaluaciones realizadas por los profesores, la naturaleza de las tareas asignadas, la falta de comprensión de los temas tratados en clase, la participación durante las lecciones y el tiempo limitado para completar las actividades (Barraza, 2008; Bedoya-Lau *et al.*, 2014).

Incluso en investigaciones recientes, como la llevada a cabo por Molina *et al.* (2023), se observó que más del 65,8 % de los estudiantes de nivel superior analizados, experimentaron notables cambios en su nivel de disfrute durante el confinamiento provocado por la pandemia de Covid-19. Estos cambios generaron diversas repercusiones en cuanto al estado de ánimo, bienestar y niveles de estrés de los estudiantes. Entre los hallazgos más significativos, se destacó que la claridad en el autoconcepto, la autorrealización y el disfrute estaban asociados a una disminución de los síntomas de estrés. Estos resultados sugieren que factores como la percepción de uno mismo y la capacidad de disfrutar pueden tener un impacto positivo en la gestión del estrés durante situaciones de confinamiento como la experimentada durante la pandemia.

Por otro lado, podría suceder que el estrés y la ansiedad generen la aparición de manifestaciones psicósomáticas, vistas como la frecuencia de aparición de síntomas tales como reacciones físicas, psicológicas o emocionales. Cuando una persona experimenta estrés intenso, ansiedad, o conflictos emocionales, su cuerpo puede responder manifestando síntomas físicos que no se pueden atribuir a una enfermedad o lesión física identificable (Eaves *et al.*, 2004; Rodríguez De Avila *et al.*, 2021; Quiceno *et al.*, 2022). Estos síntomas varían ampliamente y pueden incluir dolores de cabeza, problemas gastrointestinales, fatiga crónica, palpitaciones cardíacas, entre otros. En el estudio realizado por Bedoya-Lau *et al.* (2014), se evaluaron los niveles de estrés, las manifestaciones psicósomáticas y las estrategias de afrontamiento entre estudiantes de medicina de primer, cuarto y séptimo año en una universidad privada en Lima durante 2012. El estudio reveló que el 77,54 % de los estudiantes experimentaban estrés académico, siendo los alumnos de séptimo año quienes presentaron los niveles más altos. Además, se observó que los hombres mostraron menos estrés en comparación con las mujeres. La principal fuente de estrés identificada fue la sobrecarga de tareas y evaluaciones por parte de los profesores.

En este sentido, el diseño y la implementación de instrumentos psicométricos ha desempeñado un papel clave en la evaluación de los factores asociados al estrés académico. Entre estas herramientas se encuentra el Inventario Sistemático Cognoscitivista para el Estudio del Estrés Académico (Inventario SISCO), desarrollado por Barraza (2007), cuya finalidad es analizar el estrés académico bajo un enfoque sistémico-cognoscitivista. Este instrumento se basa en una estructura en tres dimensiones principales: las fuentes que generan estrés, las manifestaciones que provoca y las estrategias de afrontamiento empleadas por los estudiantes. En las reacciones físicas, se evalúan problemas como trastornos del sueño, la fatiga crónica, los dolores de cabeza, las dificultades digestivas, el hábito de morderse las uñas y la somnolencia excesiva. En las reacciones psicológicas se evalúan dificultades para relajarse, sentimientos de depresión y tristeza, ansiedad, angustia o desesperación, dificultades de concentración y un aumento de la irritabilidad. En lo que respecta a las reacciones conductuales, se analiza el conflicto y la tendencia a discutir, el aislamiento social, la falta de motivación para las tareas académicas y las variaciones en el consumo de alimentos. De esta manera, el Inventario SISCO se presenta como una herramienta válida y confiable para medir el estrés académico, ofreciendo una visión integral de este fenómeno desde una perspectiva psicométrica sólida.

1.3. Impacto del aprendizaje automático en la Psicología y Psiquiatría

La creciente adopción de métodos de aprendizaje automático está transformando la psicología y la psiquiatría (Chikersal *et al.*, 2021; De Lacy *et al.* 2023; Kilaskar *et al.* 2022; Pinto & Parente, 2024; Shatte *et al.*, 2019; Wani *et al.*, 2023). Estos métodos, especialmente el aprendizaje supervisado, son altamente eficaces para analizar grandes volúmenes de datos y realizar predicciones precisas (Chikersal *et al.*, 2021; De Lacy *et al.* 2023; Mentis *et al.*, 2023; Zhang, 2024). En psicología, el aprendizaje supervisado se emplea para identificar patrones en comportamientos, lo cual mejora el diagnóstico y tratamiento de trastornos (De Lacy *et al.* 2023; Fumero & Navarrete, 2016; Hawes *et al.*, 2023; Nickson *et al.*, 2023; Zhang, 2024). En psiquiatría, permite una evaluación más detallada de los síntomas y contribuye la creación de tratamientos personalizados (De Lacy *et al.* 2023; Durstewitz *et al.*, 2019). La capacidad de estos métodos para manejar y analizar datos a gran escala está revolucionando la forma en que se aborda la salud mental, ofreciendo nuevas posibilidades para mejorar la atención y la investigación en estos campos (Amador *et al.*, 2017; Chikersal *et al.*, 2021; De Lacy *et al.* 2023; Pinto & Parente, 2024; Pocco, 2022; Wani *et al.*, 2023; Zhang, 2024).

La utilización de algoritmos de clasificación para investigar la relación entre una variable de salida relacionada con la depresión y otras variables predictoras ofrece varias ventajas y una importancia significativa en el contexto de la investigación en salud mental y psiquiatría. Por ejemplo, estos algoritmos pueden identificar las variables predictoras que tienen una influencia sustancial en la probabilidad de que una persona experimente sentimientos de depresión, esto es esencial para comprender mejor los factores que pueden contribuir al desarrollo de la enfermedad. Adicionalmente, los modelos de clasificación pueden ser útiles para predecir el riesgo de depresión en individuos, basándose en sus características personales y antecedentes, lo que permitiría una identificación temprana de personas en riesgo y la implementación de medidas preventivas o intervenciones oportunas para reducir la carga de la enfermedad. Además, podrían ayudar a determinar qué tratamientos o intervenciones son más adecuados para individuos con características específicas (Chikersal *et al.*, 2021; De Lacy *et al.* 2023; Hawes *et al.*, 2023; Kilaskar *et al.* 2022; Nickson *et al.*, 2023; Pinto & Parente, 2024; Zhang, 2024).

Por ejemplo, Mentis *et al.* (2023) enfatizaron en la importancia de detectar y gestionar el estrés crónico de manera temprana como medida preventiva para diversas enfermedades, ellos resaltaron que la inteligencia artificial (IA)

y el aprendizaje automático (ML) han desencadenado un cambio significativo de enfoques en múltiples aspectos de la biomedicina, abarcando aspectos como el diagnóstico, la monitorización y la predicción de enfermedades. En ese sentido, mostraron varios modelos de aprendizaje automático para predecir el estrés, destacando su capacidad para manejar datos complejos y de alta dimensionalidad. Métodos de aprendizaje profundo, como redes neuronales profundas (DNN), logran más del 80 % de precisión en datos de resonancia magnética (MRI), mientras que bosques aleatorios (RF) y máquinas de soporte vectorial (SVM) destacan en el análisis de parámetros estadísticos derivados de datos numéricos. RF y SVM resultaron más efectivos para datos como edad, frecuencia cardíaca e historial médico, alcanzando más del 70 % de precisión, esto demuestra que tanto la inteligencia artificial (IA) como el aprendizaje automático (ML) son capaces de prever niveles significativos de estrés y diferenciar entre estados cerebrales normales y anormales, especialmente en el contexto del trastorno de estrés postraumático. En un estudio previo, [Kashani et al. \(1996\)](#) investigaron factores cruciales vinculados con la desesperanza en adolescentes a partir de una combinación de diagnósticos psiquiátricos, rasgos de personalidad y variables relacionadas con el apoyo familiar y social. Para abordar esta cuestión, seleccionaron un subconjunto de variables clave que se integraron en un modelo de red neuronal de retro-propagación, que posteriormente fue sometido a un análisis estadístico y de sensibilidad.

En el mismo sentido, [Fumero & Navarrete \(2016\)](#) demostraron la utilidad de las redes neuronales artificiales para simplificar el estudio de síntomas psicopatológicos. Utilizaron una muestra de 81 estudiantes que completaron pruebas de personalidad y crearon cuatro redes neuronales diferentes. Estas redes emplearon rasgos de personalidad como predictores de síntomas como somatización, ansiedad, depresión y disfunción social. Los resultados mostraron que esta técnica es efectiva para identificar predictores específicos de diferentes tipos de síntomas psicopatológicos. Por su parte, [Durstewitz et al. \(2019\)](#), destacaron el potencial de las redes neuronales profundas y recurrentes, exploraron además su relación con la estadística, mostrando cómo estas redes se aplican o podrían aplicarse en la psiquiatría, brindando oportunidades para la predicción y clasificación de trastornos psiquiátricos. Además, plantearon la posibilidad de combinar modelos computacionales interpretables semánticamente con técnicas de aprendizaje automático estadístico para obtener información sobre los mecanismos neuronales y conductuales.

Adicionalmente, [Chikersal et al. \(2021\)](#) presentaron un enfoque novedoso para la detección y predicción de la depresión, basado en el análisis de síntomas a lo largo del tiempo. Utilizaron dos algoritmos de aprendizaje automático: regresión logística y un clasificador de refuerzo de gradiente (Gradient Boosting), los cuales fueron combinados mediante el uso de un clasificador en conjunto (AdaBoost) para optimizar los resultados. El modelo mostró una precisión del 85,7 % al identificar la depresión posterior al semestre y del 81,3 % al predecirla hasta 11 semanas antes de su aparición. Estos hallazgos tienen el potencial de transformar las intervenciones tecnológicas en tiempo real, mejorando así el diagnóstico, tratamiento y prevención de la depresión.

Por otro lado, [De Lacy et al. \(2023\)](#) aplicaron aprendizaje automático utilizando redes neuronales artificiales (ANN), XGBoost y regresión logística con ElasticNet, integrados en una arquitectura de aprendizaje evolutivo para optimizar la selección de características y los hiperparámetros. Con 160 predictores multidimensionales, estos modelos lograron una $AUC \geq 0,94$ en datos no vistos. Los algoritmos ANN y XGBoost superaron significativamente a ElasticNet, demostrando su capacidad para modelar relaciones no lineales en patologías psiquiátricas. Además, los predictores psicosociales fueron más efectivos que las métricas neuronales, subrayando el carácter multifactorial de estas afecciones. Los resultados mostraron modelos robustos para predecir la ansiedad y depresión.

1.4. Nuestro aporte en la aplicación de machine learning para identificar factores de ansiedad, estrés y depresión en adolescentes

En esta investigación tuvo como objetivo identificar factores determinantes relacionados con la ansiedad, el estrés y la depresión en adolescentes, con el fin de brindar apoyo psicológico y emocional que favorezca un sano desarrollo, promoviendo una alta autoestima, motivación y confianza en los estudiantes. Para ello, se aplicaron diferentes algoritmos de Machine Learning para clasificación, debido a que la variable objetivo es de respuesta binaria (sí/no). Los algoritmos utilizados fueron: Regresión Logística, k-NN, Árbol C5.0, Árbol CART, Bagging, Random Forest, Gradient Boosting Machine (GBM), XGBoosting, Red Neuronal Perceptrón, Máquina de Soporte Vectorial con kernel lineal (SVL) y Máquina de Soporte Vectorial con kernel radial (SVM).

Adicionalmente, se aplicaron métodos de ensamble, los cuales, en el contexto del Machine Learning y la Estadística, son técnicas que combinan

múltiples modelos o algoritmos de predicción para obtener un rendimiento más robusto y preciso que el que se podría lograr con un solo modelo. Estos métodos se basan en la premisa de que la combinación de múltiples modelos puede reducir el sesgo y la variabilidad inherentes a cada uno, lo que puede mejorar la capacidad de generalización y la precisión de las predicciones. Son ampliamente utilizados en estos contextos por su capacidad para mejorar la precisión de las predicciones y reducir el sobreajuste. Algunos de los métodos de ensamble más conocidos incluyen algoritmos *Bagging* (*Bootstrap Aggregating*) como el Random Forest y algoritmos *Boosting* como el AdaBoost y Gradient Boosting (Dixit, 2017). Cabe mencionar la investigación realizada por Gamboa & Salinas (2022), quienes emplearon algoritmos de Machine Learning y de ensamble con el objetivo de identificar los principales predictores de la situación académica de un estudiante universitario luego de que transcurrieron seis semestres desde su ingreso.

Previamente, se realizó una selección de variables debido a que es problema común en algoritmos de Machine Learning; esto sucede cuando se tiene un gran número de variables predictoras, pero solo unas pocas de ellas realmente influyen en la variable de respuesta que se desea predecir; es decir, un problema de alta dimensionalidad. En este contexto, uno de los principales desafíos es realizar una selección de variables para identificar las covariables que no tienen influencia real en la respuesta (variable target) y eliminarlas del modelo de predicción. Tal como lo menciona Aneiros *et al.* (2015), de lo que se trata es de encontrar un equilibrio entre flexibilidad y complejidad en el modelo, al tiempo que se identifican y eliminan las variables irrelevantes para mejorar la precisión de la predicción en situaciones de alta dimensionalidad. Finalmente, se darán algunas reflexiones acerca de las ventajas de los algoritmos de Machine Learning aplicados en el área de la psicología.

2. MÉTODO

2.1. Diseño empleado

Inicialmente se realizó una etapa observacional por parte de la docente orientadora de la Institución Educativa del Departamento del Tolima, esta actividad permitió identificar factores que afectan el bienestar emocional de los estudiantes, tales como conductas agresivas, problemas familiares, uso no supervisado de tecnología, calificaciones bajas, desánimo hacia la realización de tareas, así como una tendencia a no asumir las consecuencias de sus acciones. Además, el aislamiento causado por la pandemia del COVID-19 hacía suponer el aumento de estos factores, resultando en cambios emocio-

nales complejos en los adolescentes, como alteraciones en sus relaciones interpersonales, falta de motivación, depresión o ansiedad. Debido a lo anterior, se elaboró un instrumento el cual fue compartido con los docentes, directivos y administrativos de la Institución Educativa con el fin de realizar la sensibilización y compartir opiniones acerca del alcance del proyecto.

2.2. Instrumentos para la recolección de la información

El cuestionario aplicado fue construido teniendo como base el Inventario Sistemático Cognoscitivista para el estudio del estrés académico (instrumento SISCO), cuyo constructo teórico se basa en la clasificación de Rossi (2001), ampliando las posibilidades de estudiar este constructo con otras preguntas relacionadas a los objetivos y el contexto de esta investigación (Barraza, 2007; Bedoya-Lau *et al.*, 2014).

El instrumento se dividió en grupos de preguntas. El primer grupo fue de caracterización, en el que se incluyó el género, edad, grado que cursaba actualmente y el grupo al que correspondía el respectivo grado. El segundo grupo observó las manifestaciones psicósomáticas (reacciones o síntomas físicos, psicológicos y comportamentales ante el estrés). En el tercer grupo se evaluaron las características estresoras en el ámbito académico. Por último, se evaluó una variable binaria relacionada con que si la pandemia del COVID-19 había influenciado en el aumento del estrés, la ansiedad o la depresión, y una variable binaria relacionada con la depresión.

Se usó el coeficiente Alfa de Cronbach cuyo valor mide la confiabilidad o consistencia interna de un conjunto de preguntas o ítems en una escala de medición (Cronbach, 1951). De esta manera se calculó la confiabilidad mediante la consistencia interna y la determinación del coeficiente cuyo resultado fue de 0,93.

El instrumento fue aplicado presencialmente de manera voluntaria, anónima e individual luego de realizar la sensibilización y socialización del proyecto. De esta manera, esta propuesta sigue los lineamientos éticos que reflejan la responsabilidad social implicada en cualquier investigación, respetando la dignidad humana, el derecho al buen nombre de los participantes, el manejo adecuado y cuidadoso de los datos, así como la confiabilidad y veracidad del proceso de investigación. Con respecto a la confidencialidad de la información, se garantizó en todo momento la privacidad y el buen nombre de

los actores participantes, sin publicar nombres ni identificaciones reales a aquellos que conozcan los resultados de la investigación.

2.3. Participantes

La Institución Educativa es de población mixta con única jornada académica durante la mañana. El instrumento fue aplicado a 553 estudiantes de básica secundaria y media (de sexto a undécimo grado). Cabe resaltar que la encuesta se realizó de forma presencial, porque el acceso a internet en la Institución Educativa es limitado. Además, como el objetivo era impactar a la mayoría de los estudiantes se eligió un muestreo accidental no probabilístico, con el fin de recabar las opiniones de los participantes de manera ágil. Este tipo de muestreo se enmarca en la categoría no probabilística, ya que la obtención de la muestra no siguió un proceso aleatorio. La naturaleza accidental de esta muestra reside en el hecho de que la selección de los individuos se realizó de manera en que estuvieran disponibles y accesibles en el lugar y momento de la aplicación del instrumento, sin la aplicación de criterios predefinidos. Dicho de otra manera, los individuos incluidos en la muestra fueron resultado de circunstancias accidentales, sin un plan preestablecido de selección (Otzen & Manterola, 2017).

La edad promedio de los participantes fue 14,05 años (D.E.=1,92), siendo 11 la edad mínima y 20 la máxima, las edades de 18, 19 y 20 fueron las de menor frecuencia. Respecto a la identidad de género, 231 (49,1 %) son del género masculino, 220 (46,8 %) del femenino, 6 (1,3 %) se clasificaron en no binarios y 13 (2,8 %) prefirieron no revelar su identidad. El grado con menor cantidad de instrumentos aplicados fue el grado once con 51, seguido del grado sexto con 77, en los grados séptimo, octavo, noveno y décimo se aplicaron 84, 89, 81 y 88 respectivamente.

2.4. Procedimientos desarrollados

Después de la recolección de los datos, se llevó a cabo un pre-procesamiento de los mismos para el proceso de modelado. Durante esta etapa, se excluyeron los individuos con datos incompletos, lo que dejó un total de 470 observaciones completas. Se efectuó un análisis descriptivo univariado y bivariado utilizando tanto las variables predictoras como la variable dependiente a predecir (depresión).

Para seleccionar las variables predictoras más relevantes, existen algoritmos de Filtro, Wrapper (RFE y BORUTA) y Embedded (Ridge y Lasso). Se aplicó el

algoritmo BORUTA porque este algoritmo es robusto y superior en la identificación de características relevantes debido a su capacidad para manejar relaciones complejas y no lineales (Chandrashekar & Sahin, 2014).

El algoritmo BORUTA duplica el conjunto de variables, mezcla los valores dentro de cada una y crea lo que se denomina *variables sombra* o *shadow variables*. Luego, un clasificador es entrenado utilizando el algoritmo *Random Forest* en este conjunto de datos, y se calcula el *Mean Decrease Accuracy* o el *Mean Decrease Impurity* para cada variable. Cuanto mayor es el puntaje obtenido, más relevante es la variable en cuestión. En resumen, el algoritmo *BORUTA* intenta validar la importancia de una variable comparándola con copias mezcladas aleatoriamente de la misma (Lantz, 2019; Gamboa & Salinas, 2022).

El conjunto de datos se dividió aleatoriamente, asignando el 80 % de los registros para entrenamiento y el 20 % restante para evaluación. Además, se comprobó que ambas particiones mantuvieran una proporción de observaciones similar a la distribución original de las categorías de la variable objetivo: depresión (con y sin depresión). Las variables numéricas fueron normalizadas y las categóricas fueron convertidas en variables dummy. Al tener un 76,17 % de estudiantes con depresión y un 23,83 % sin depresión, no se consideró necesario realizar un balanceo de la data de entrenamiento ya que la proporción es representativa de la muestra total y refleja la realidad observada en el contexto del estudio. Además, la clase minoritaria cuenta con suficientes observaciones para permitir un aprendizaje efectivo del modelo sin necesidad de aplicar técnicas de balanceo.

2.5. Análisis de datos

Para la fase de modelado, se empleó la validación cruzada de 10 capas para estimar y seleccionar los hiperparámetros de los algoritmos que se entrenaron (James *et al.*, 2013). Los algoritmos utilizados fueron: Regresión Logística, k-NN, Árbol C5.0, Árbol CART, Bagging, Random Forest, Gradient Boosting Machine (GBM), XGBoosting, Red Neuronal Perceptrón, Máquina de Soporte Vectorial con kernel lineal (SVL) y Máquina de Soporte Vectorial con kernel radial (SVM). Esto se realizó para seleccionar los mejores algoritmos en función a su capacidad predictiva (Gamboa & Salinas, 2022). En la Tabla 1 se presentan los hiperparámetros tuneados con la finalidad de obtener el mayor accuracy en la muestra de entrenamiento con validación cruzada.

Tabla 1.

Hiperparámetros de cada uno de los algoritmos entrenados

ALGORITMO	HIPERPARÁMETRO(S) TUNEADO(S)	HIPERPARÁMETRO(S) SELECCIONADO(S)	ACCURACY EN EL ENTRENAMIENTO CON VALIDACIÓN CRUZADA
Regresión Logística	No tiene hiperparámetros a entrenar		0,7717639
k-nn	k con valores de 5, 7, 9, 11 y 13	k = 13	0,7902560
Árbol C5.0	Algoritmo basado en árbol o en reglas, boosting con 1, 10, 20, 30, 40, 50, 60, 70, 80 y 90 iteraciones	Algoritmo basado en reglas con 20 iteraciones en el boosting	0,7981508
Árbol CART	10 valores del parámetro de complejidad (cp) entre 0 y 0,05 igualmente espaciados	cp = 0,037	0,8090327
Bagging	50 y 100 árboles	100 árboles	0,7981508
Random Forest	Se tunearon los valores 2, 7, 12,17 y 23 para el hiperparámetro del número de variables a seleccionar en el entrenamiento de cada nodo	Número de variables = 17	0,8219061
Gradient Boosting Machine (GBM)	Profundidad del árbol desde 1 hasta nivel 10, learning rate de 0 hasta 0,5 y número de árboles entre 50 a 500 árboles.	Profundidad del árbol de 3, learning rate de 0,1 y número de árboles de 50	0,8061166
XGBoosting	Profundidad del árbol con valores de 1, 2 y 3, muestra de variables de 0,6 y 0,8, muestra de individuos de 0,5, 0,75 y 1, tasa de aprendizaje de 0,3 y 0,4	Profundidad del árbol de 2, muestra de variables de 0,6, muestra de individuos de 0,5, tasa de aprendizaje de 0,4	0,7928876
Red Neuronal Perceptrón	Capa oculta con 1, 3, 5, ..., 17, 19 neuronas y un decay entre 0 y 0,05	Capa oculta con 19 neuronas y un decay de 0,1	0,7955192
Máquina de Soporte Vectorial con kernel lineal	Costo con 10 valores entre 0 y 128	Costo de 1	0,7849929
Máquina de Soporte Vectorial con kernel radial	Costo con 10 valores entre 0 y 128 y un sigma con 10 valores entre 0 y 0,05	Costo de 0,5 y sigma de 0,03	0,7981508

Luego, se creó un modelo de ensamble utilizando los tres algoritmos que ofrecieron los mejores resultados en el entrenamiento y que no estaban correlacionados entre sí. Este modelo se basó en el promedio de las probabilidades obtenidas, y los indicadores se optimizaron ajustando el punto de corte según lo sugerido por la curva ROC. Los métodos de ensamble combinan varios algoritmos de aprendizaje para lograr un rendimiento superior. Entre estos métodos se encuentran el promedio simple, el promedio ponderado y el voto mayoritario, tal como se describe en los trabajos de [Alfaro \(2019\)](#) y [Dixit \(2017\)](#). Para mantener la evaluación sin balancear, se emplearon métricas robustas a la desproporción de las clases, como la sensibilidad, la especificidad y el accuracy balanceado.

A través del entrenamiento de los algoritmos se buscó proporcionar información valiosa para la identificación temprana y el abordaje de problemas emocionales en los adolescentes a fin de que el especialista en psicología

realice un plan de tratamiento que se ajuste mejor a las necesidades de cada uno. Cabe aclarar que las respuestas binarias a una única pregunta pueden ser indicativas de sentimientos de tristeza, pero no necesariamente de diagnóstico clínico de depresión. La interpretación de los resultados debe hacerse con sensibilidad y considerando el contexto.

2.6. Software

Para realizar el análisis de la información se empleó el software de código abierto R (version 4.2.3), desarrollado por [R core team \(2023\)](#) y el lenguaje de programación Python (versión 3), ([Van Rossum & Drake, 2009](#)), ya que proporcionan herramientas robustas para la implementación de algoritmos de clasificación, gracias a su amplia gama de paquetes, su habilidad para gestionar grandes volúmenes de datos, y su soporte para visualización y análisis estadístico. Al aplicar estos algoritmos, los usuarios pueden realizar predicciones acertadas, identificar patrones en los datos y optimizar la toma de decisiones fundamentada en datos.

3. RESULTADOS

Como resultado de la comprensión de los datos y aplicando el algoritmo BORUTA se seleccionaron las siguientes variables predictoras relacionadas a la situación de depresión o no de un estudiante: *Ansiedad, Fatiga Crónica, Falta apetito, Difícil decidir, Somnolencia, Intranquilidad, Aislamiento, Conflictivo, Falta concentración, Trastorno sueño, Actividades evaluativas, Falta motivación, Carga económica, Dolor de cabeza, No entiende clase, Agresivo, irritable, Rascar, morder uña, Género, Carácter del profesor, Poco tiempo de actividad, Tipos de actividades.*

Las variables *Covid-19, Sobrecarga escolar, Competir grupo, Problemas de digestión, Edad, Grado, Grupo y Participar en clase* fueron retiradas del modelo por tener poco poder predictivo. Posteriormente, la variable *Desmotivación escolar* fue retirada al estar altamente correlacionada con las demás variables predictoras.

Al preguntarle a los estudiantes sobre si alguna vez habían tenido sentimientos persistentes de tristeza, sensación de vacío o pérdida de interés en las actividades que realiza cotidianamente (“depresión”), se pudo observar que, de los 470 estudiantes, el 76,2 % manifestaron que sí habían presentado estos sentimientos y el 23,8 % que no (ver Figura 1). Estos resultados podrían ser un indicio de que se debe trabajar con los estudiantes con el fin de contrarrestar o minimizar dichos -sentimientos.

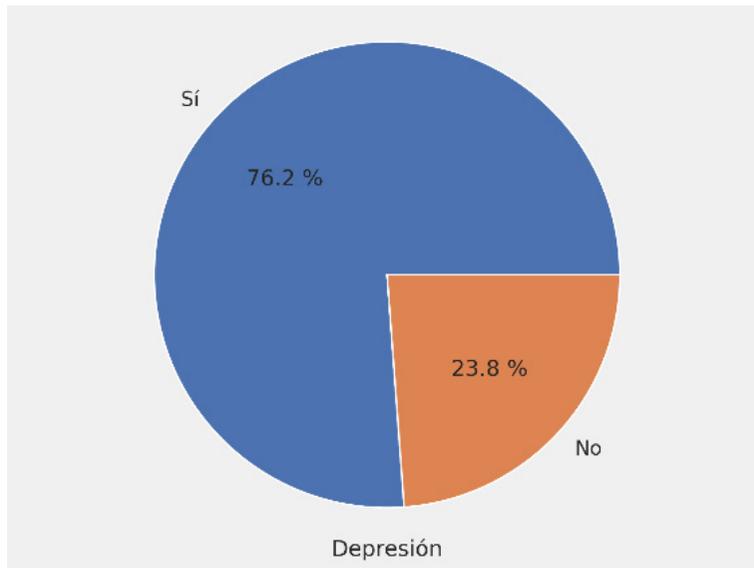


Figura 1. Distribución de los Estudiantes según la Variable Target Depresión

Con relación a las variables predictoras numéricas seleccionadas, se puede apreciar en la Figura 2 que la calificación promedio estuvo entre un mínimo de 2,13 (conflictivo) y un máximo de 3,15 (rascar y morder uña).

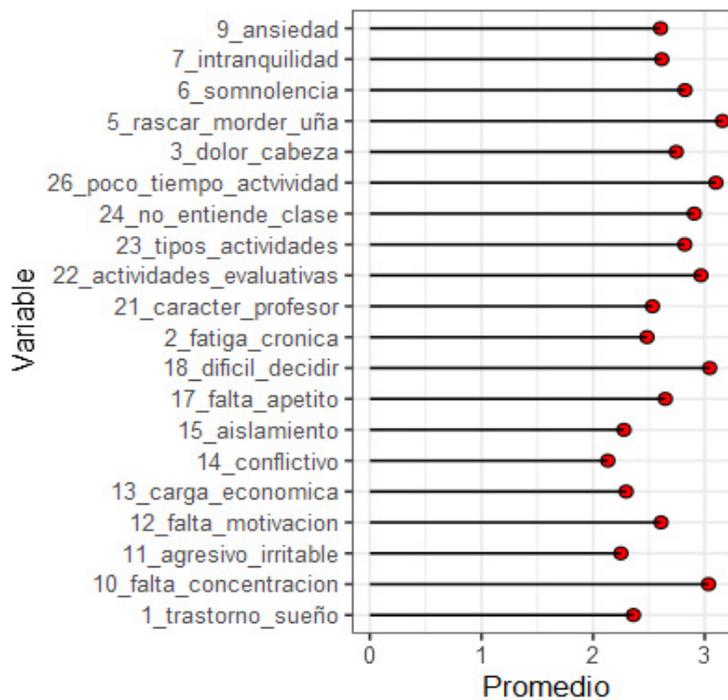


Figura 2. Calificación Promedio por Variable Predictor Numérica

También se puede apreciar en la Figura 3 el poder predictivo de todas las variables predictoras seleccionadas por el algoritmo BORUTA, mostrando promedios bajos para los estudiantes que no tuvieron depresión y promedios altos para los estudiantes que manifestaron si tener depresión. En todos los casos se puede observar que los promedios para ambas categorías se alejan del promedio general para todas las variables predictoras.

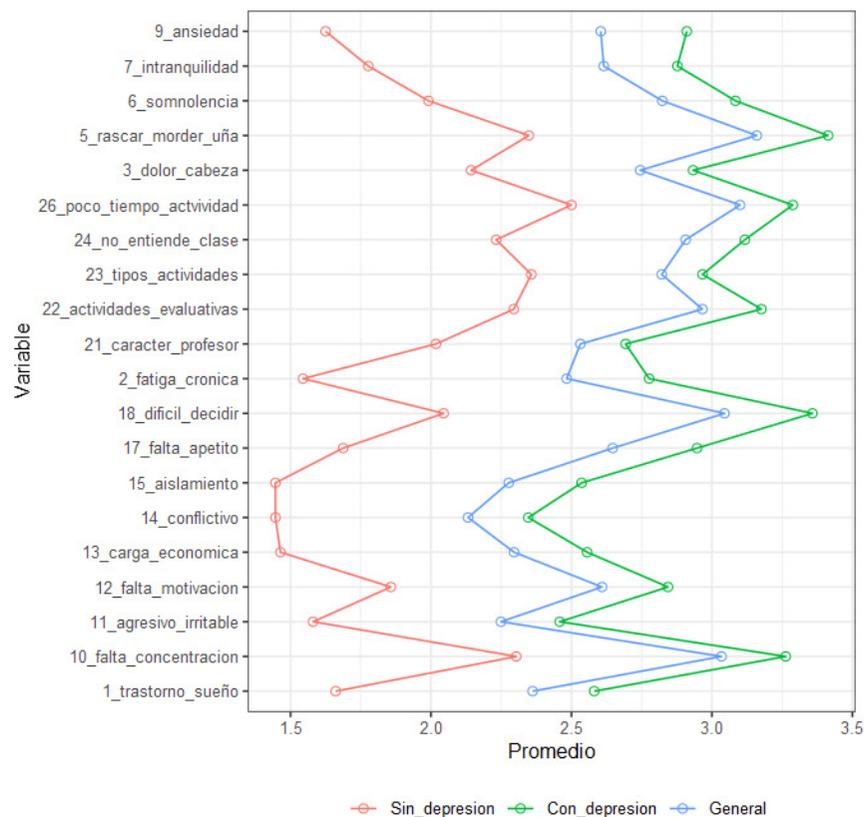


Figura 3. Calificación Promedio por Variable Predictor Numérica Vs la Variable Target Depresión.

En la Figura 4 se puede observar que los algoritmos que son más estables son el Bagging, el árbol C5 y el k-NN. Estos tres algoritmos fueron ensamblados. Los algoritmos de: Regresión Logística, Redes Neuronal Perceptrón Multicapa, Máquina de Soporte Vectorial con kernel radial, Random Forest, GBM y XGBoosting no fueron ensamblados al presentar resultados altamente correlacionados.

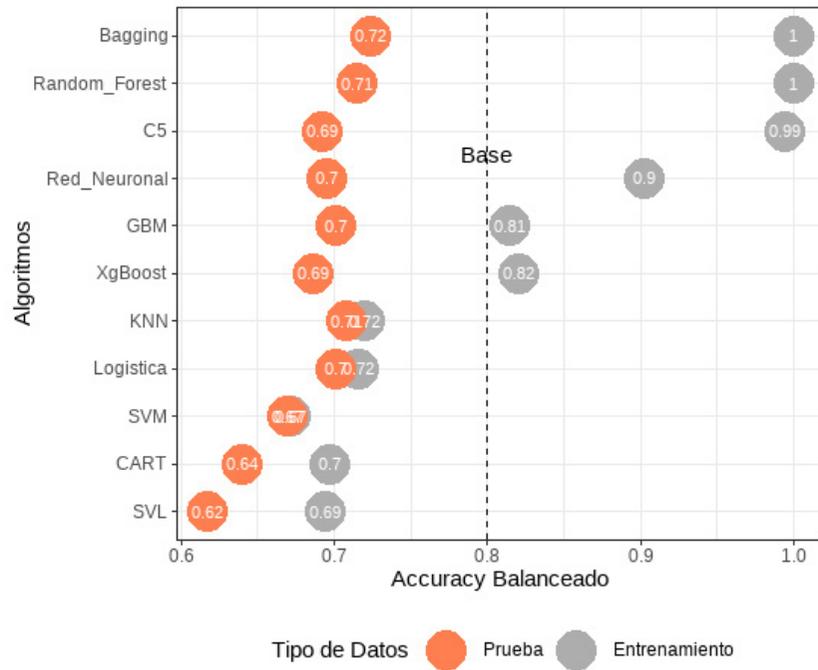


Figura 4. Comparación del Accuracy Balanceado para los 12 Algoritmos de Machine Learning

Respecto al ensamble de los modelos, los resultados de la sensibilidad, especificidad y accuracy balanceado para cada uno de los tres modelos elegidos se muestran en la Tabla 2, así también para el ensamble de éstos, usando el punto de corte tradicional (umbral de 0,5) y el óptimo sugerido por la curva ROC que se aprecia en la Figura 5. Con este valor óptimo (umbral de 0,662), la probabilidad de detectar correctamente a los estudiantes con depresión es de 0,8450704, mientras que la detección de estudiantes que no manifiestan tener depresión se realiza con una probabilidad de 0,86364.

Tabla 2.

Comparación de indicadores para los tres algoritmos y el ensamble con y sin punto de corte óptimo.

ALGORITMO	SENSIBILIDAD	ESPECIFICIDAD	ACCURACY BALANCEADO
Ensamble con umbral óptimo de 0,662	0,8450704	0,8636364	0,8543534
Bagging (con número de árboles de 100)	0,9014085	0,5454545	0,7234315
Ensamble con umbral de 0,5	0,9295775	0,5000000	0,7147887
k-NN (con k de 13)	0,9154930	0,5000000	0,7077465
Árbol C5	0,9295775	0,4545455	0,6920615

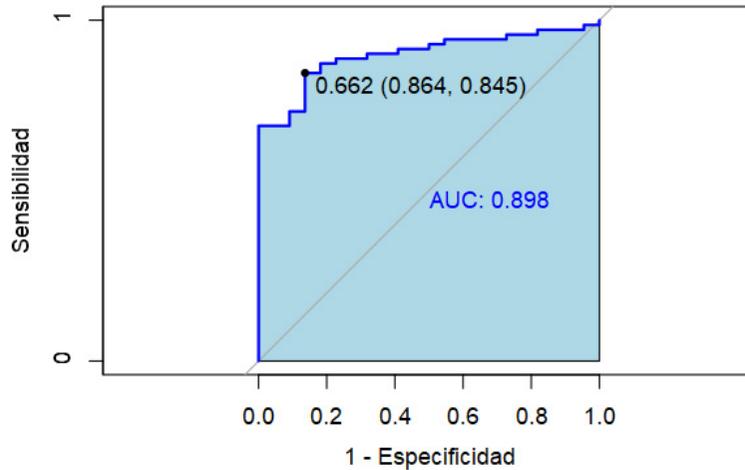


Figura 5. Curva ROC para el Ensamble con los tres Modelos

La aplicación de los modelos anteriores, permitieron encontrar las principales variables que mejor discriminan a un estudiante que contestó afirmativamente tener sentimientos persistentes de tristeza, sensación de vacío o pérdida de interés en las actividades que realiza cotidianamente (*depresión*) de uno que no lo tiene; estas son: *fatiga crónica, difícil decidir, la ansiedad, falta de apetito, conflictivo, actividades evaluativas y somnolencia* (Ver Figura 6).

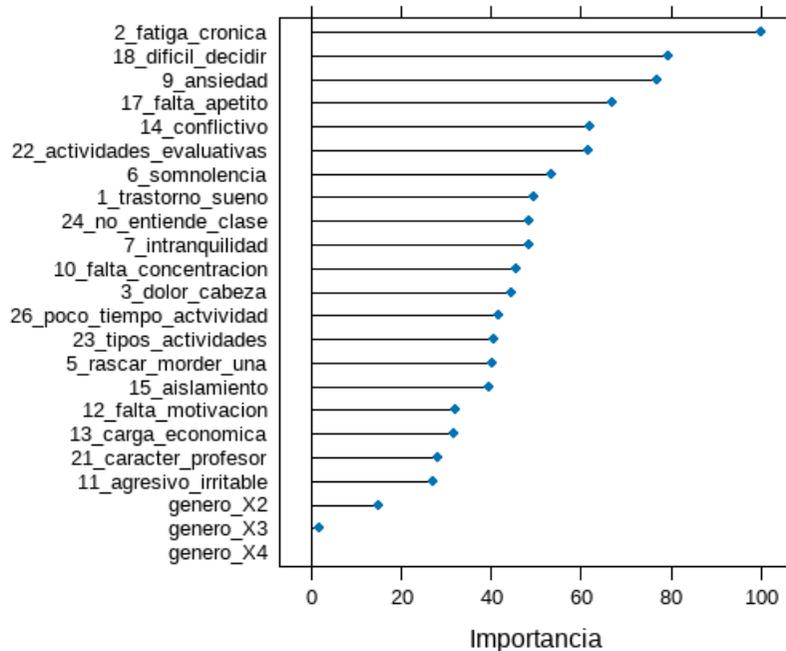


Figura 6. Importancia de las Variables Predictoras

4. DISCUSIONES

El objetivo de este trabajo fue identificar factores determinantes relacionados a la ansiedad, estrés y depresión en adolescentes con el fin de brindar apoyo psicológico y emocional generando un sano desarrollo, con alta autoestima, motivación y confianza en ellos mismos.

Teniendo en mente dicho propósito, se mostró que es esencial detectar y gestionar temas como ansiedad y el estrés de manera temprana como medida preventiva para diversas enfermedades tales como la depresión. En ese sentido, los algoritmos de Machine Learning han mostrado buen desempeño y han permitido seleccionar las variables independientes que más aportan a la variable target; así mismo, permiten manejar grandes volúmenes de datos de manera eficiente y puede ser replicado en otros contextos para identificar factores de riesgo similares en diferentes poblaciones (Chikersal *et al.*, 2021; De Lacy *et al.* 2023; Durstewitz *et al.*, 2019; Fumero & Navarrete, 2016; Kashani *et al.*, 1996; Mentis *et al.*, 2023; Nickson *et al.*, 2023; Pinto & Parente, 2024; Wani *et al.*, 2023; Zhang, 2024), incluso mostrando las fortalezas de los métodos de ensamble para establecer un balance entre el sesgo y la variabilidad, mejorando la capacidad de generalización y la precisión de las predicciones (Gamboa & Salinas, 2022; Nickson *et al.*, 2023).

Aunque la revisión llevada a cabo por Nickson *et al.* (2023) señala el uso de diversos métodos de aprendizaje automático supervisado para la clasificación de la depresión, incluyendo modelos de regresión (en 13 estudios), Random Forest (en 8 estudios), XGBoost (en 8 estudios) y SVM (en 7 estudios), no se hace mención del uso de técnicas de ensamblaje de modelos. En contraste, nuestro estudio demostró que el empleo de esta técnica de ensamble mejoró sustancialmente la precisión de las predicciones, tal como sucedió en el trabajo realizado por Gamboa & Salinas (2022).

Utilizando el umbral óptimo de 0,662 sugerido por la curva ROC (con un área bajo la curva AUC de 0,898) en el modelo ensamblado, se lograron mejores indicadores de sensibilidad y especificidad en comparación con el umbral tradicional de 0,5. Este punto de corte de 0,662 permitió identificar correctamente a los estudiantes con depresión con una probabilidad de 0,845, y a los estudiantes que no presentan síntomas de depresión con una probabilidad de 0,864. El valor de AUC obtenido en este estudio es excelente, ya que se encuentra cerca del 0,9. De hecho, en la revisión realizada por Nickson *et al.* (2023) sobre la literatura existente en el uso de métodos de aprendizaje automático combinados con registros médicos electrónicos para la

predicción de la depresión, se encontraron valores entre 0,7 a 0,9, con un promedio de 0,78.

La interpretabilidad y la visualización son fundamentales para una implementación eficaz de los algoritmos de aprendizaje automático, como indican [Nickson et al. \(2023\)](#) y [De Lacy et al. \(2023\)](#). Es crucial identificar y detallar los predictores específicos que afectan la variable objetivo. En nuestro estudio, hemos identificado las variables que mejor discriminan entre estudiantes que reportan sentimientos persistentes de tristeza, vacío o pérdida de interés en sus actividades diarias (indicativos de depresión) y aquellos que no presentan estos síntomas. Las variables más significativas son: fatiga crónica, dificultad para tomar decisiones, ansiedad, falta de apetito, comportamiento conflictivo, carga de actividades evaluativas y somnolencia. Muchas de estas variables también fueron destacadas en la revisión de [Nickson et al. \(2023\)](#). Además, utilizamos el algoritmo BORUTA ([Gamboa & Salinas, 2022](#); [Lantz, 2019](#)) para visualizar el poder predictivo de cada variable.

A nivel psicosocial, el estudio reveló aquellos componentes que los estudiantes enfrentan en los ámbitos emocional, social y físico. Estos factores los llevan diariamente a experimentar episodios de ansiedad, fatiga crónica, pérdida de apetito, dificultad para tomar decisiones, somnolencia, intranquilidad, aislamiento, comportamiento conflictivo, falta de concentración y trastornos del sueño. Coincidiendo de cierta manera con los estudios desarrollados por [Bedoya-Lau et al. \(2014\)](#), [Lopez-Vega et al. \(2020\)](#) y [De Lacy et al. \(2023\)](#).

Estas situaciones son determinantes y, si no se lleva a cabo un proceso de prevención oportuno, pueden conducir a la depresión, convirtiéndose en un problema de salud mental. Es por ello, que es crucial estar atentos a los cambios que pueda experimentar el adolescente en su forma de pensar, sentir y comportarse. De esta manera, se resalta la importancia de este estudio como un insumo invaluable para identificar los ítems más relevantes, esto permitirá ofrecer a los adolescentes mecanismos de protección a través de talleres lúdico-pedagógicos que fomenten la motivación y la participación activa en clase. Este tipo de talleres podrían brindar la oportunidad de desarrollar estrategias prácticas para mitigar el impacto progresivo de la depresión en los estudiantes ([Bustos & Russo, 2018](#); [Lopez-Vega et al., 2020](#)). Incluso, se podría pensar en el desarrollo de herramientas de monitoreo en tiempo real basadas en los algoritmos que detecten signos tempranos de depresión, permitiendo intervenciones más rápidas y efectivas.

En ese sentido, se subraya la relevancia de fomentar programas y proyectos que promuevan la salud mental de los adolescentes, incluso desde la infancia, que brinden satisfacción y logro personal, dado su impacto en las esferas psicológicas y académicas, así como su función crucial en la mejora del bienestar. Estos programas o herramientas deben orientarse a fortalecer la red de servicios de salud y fomentar estilos de vida saludables que respondan a las necesidades concretas del entorno. Asimismo, deben contribuir a minimizar el impacto adverso de los problemas psicosociales en diversas comunidades, instituciones educativas, organizaciones y familias (Bustos & Russo, 2018; González-Hernández *et al.*, 2019; Molina *et al.*, 2023; Núñez *et al.*, 2017).

Cabe recordar que, si bien estos métodos de análisis proporcionan información valiosa para la identificación temprana y el abordaje de problemas emocionales en los adolescentes, es esencial considerar el contexto específico en el que se realizó el estudio. El contexto puede incluir factores como el entorno familiar, las dinámicas escolares y las características individuales del adolescente (Alcindor-Huelva *et al.*, 2019; Amador *et al.*, 2017; Bedoya-Lau *et al.*, 2014; De Lacy *et al.*, 2023; Erazo & Martínez, 2024; Lopez-Vega *et al.*, 2020; Molina *et al.*, 2023; Núñez *et al.*, 2017; Salazar *et al.*, 2016). Estos elementos influyen en cómo los problemas emocionales se manifiestan y se manejan en la vida real. Por lo tanto, el especialista en psicología debe integrar esta comprensión contextual al desarrollar un plan de tratamiento. Un enfoque personalizado y contextualizado permite que las intervenciones sean más efectivas y se ajusten mejor a las necesidades particulares de cada individuo, mejorando así los resultados del tratamiento (Bustos & Russo, 2018; De Lacy *et al.*, 2023; Erazo & Martínez, 2024; Martínez de Velasco *et al.*, 2023; Nickson *et al.*, 2023).

5. CONCLUSIONES

El estudio proporciona una visión detallada de los factores psicosociales que afectan a los adolescentes, como episodios de ansiedad, fatiga crónica, cambios en el apetito y trastornos del sueño. Los resultados coinciden con investigaciones como la de Nickson *et al.* (2023), destacando la importancia de abordar componentes emocionales, sociales y físicos en el bienestar de los adolescentes. Es fundamental destacar la importancia de la detección temprana, el uso de herramientas de Machine Learning y la adopción de enfoques preventivos y personalizados para abordar la ansiedad, el estrés y la depresión en adolescentes. Sin embargo, es igualmente importante considerar la interpretabilidad y la visualización al implementar estos algoritmos

de aprendizaje automático, como sugieren [Nickson et al. \(2023\)](#). Identificar y describir claramente los predictores que influyen en la variable objetivo es esencial para comprender mejor los resultados. Además, es crucial evaluar los indicadores de ajuste, como la sensibilidad, especificidad y precisión (accuracy), para entender el desempeño de cada modelo aplicado. Este análisis es vital para tomar decisiones informadas basadas en los resultados obtenidos.

Se sugiere realizar talleres lúdico-pedagógicos que ofrezcan mecanismos de protección y estrategias prácticas para mitigar el impacto progresivo de la depresión en los estudiantes. Además, sería valioso desarrollar herramientas de monitoreo en tiempo real basadas en algoritmos que detecten signos tempranos de depresión, permitiendo intervenciones más rápidas y efectivas. Estas herramientas podrían integrarse con aplicaciones móviles o plataformas digitales que faciliten el seguimiento continuo del estado emocional de los adolescentes, promoviendo una comunicación abierta entre los estudiantes, sus familias y los profesionales de la salud mental. Asimismo, combinar estos enfoques tecnológicos con intervenciones personalizadas podría optimizar el apoyo emocional y psicológico, adaptándose a las necesidades individuales de cada estudiante.

6. LIMITACIONES Y RECOMENDACIONES

En primer lugar, dado que la muestra fue seleccionada de manera no aleatoria, las conclusiones alcanzadas no pueden aplicarse de manera general a todos los estudiantes de básica secundaria y media en el Departamento del Tolima, y mucho menos a nivel nacional en Colombia.

En segundo lugar, a medida que observamos que problemas familiares, el uso excesivo de medios tecnológicos, la falta de presencia parental y la carencia de autoestima contribuyen al aumento de trastornos psicológicos en los adolescentes, quienes frecuentemente enfrentan dificultades para regular positivamente sus emociones, se sugiere que futuras investigaciones amplíen su alcance considerando variables que aborden aspectos relacionados con estos problemas.

7. HIGHLIGHTS (IDEAS CLAVE)

El análisis de depresión en adolescentes de básica secundaria y media usando algoritmos de Machine Learning permitió identificar factores determinantes relacionados a la ansiedad, estrés y depresión en adolescentes de básica secundaria y media (grado sexto a once) de una Institución Educativa

del Departamento del Tolima. Se encontraron que estas son: la fatiga crónica, difícil decidir, la ansiedad, falta de apetito conflictivo, actividades evaluativas y somnolencia. La probabilidad de detectar correctamente a los estudiantes con depresión fue de 0,8450704, mientras que para los estudiantes que no manifiestan tener depresión fue 0,8636364.

Agradecimientos: A la Oficina de Proyección Social de la Universidad del Tolima y a los docentes, directivos y administrativos de la Institución Educativa.

Conflicto de interés: Ninguno.

Financiamiento: Nombre del proyecto: Modelo explicativo del estrés, la ansiedad y la depresión en jóvenes de Básica Secundaria y Media de una Institución Oficial del Departamento del Tolima. Entidad Financiadora: Oficina de Proyección Social, Universidad del Tolima. Código proyecto: BPUT-027.

Descargo de responsabilidad: Los autores del presente texto participaron tanto en el proceso investigativo como en la elaboración del manuscrito. Las opiniones aquí manifestadas son propias y producto del ejercicio de investigación realizado.

REFERENCIAS

- Alcindor-Huelva, P., Delgado-Campos, A. M., Sipos-Gálvez, L., Fernández-Úbeda, C., & Rodríguez-Solano, J. J. (2019). Acoso escolar, conductas autolesivas, ideación, e intentos autolíticos en una muestra clínica de un centro de salud mental. *Revista De Psiquiatría Infanto-Juvenil*, 36(4), 14–23. <https://doi.org/10.31766/revpsij.v36n4a3>
- Alfaro E, Gámez M, García N. (2019). *Ensemble Classification Methods with Applications in R*. New Jersey: John Wiley & Sons, Ltd.
- Amador Moncada, J. A., Granada Díaz, H. A., Redondo Ortegón, J. M. & Olivar Tos, o. (2017). Dinámicas No-Lineales y No-Suaves en Procesos Estrés-Enfermedad. *Ciencia en Desarrollo*, 8(1), 9-19.. <https://doi.org/10.19053/01217488.v8.n1.2017.5462>
- Aneiros, G., Ferraty, F., & Vieu, P. (2015). Variable selection in partial linear regression with functional covariate. *Statistics*, 49(6), 1322-1347. <https://doi.org/10.1080/02331888.2014.998675>
- Barraza, A. (2007, 26 de febrero). Inventario SISCO del estrés académico. Propiedades psicométricas. *Revista Psicología Científica.com*, 9(13). <https://pscient.net/8vd5j>
- Barraza Macías, A. (2008). El estrés académico en alumnos de maestría y sus variables moduladoras: un diseño de diferencia de grupos. *Avances En Psicología Latinoamericana*, 26(2), 270-289. <https://revistas.urosario.edu.co/index.php/apl/article/view/67>
- Bedoya-Lau, Francisco N., Matos, Luis J., & Zelaya, Elena C. (2014). Niveles de estrés académico, manifestaciones psicósomáticas y estrategias de afrontamiento en alumnos de la facultad de medicina de una universidad privada de Lima en el año 2012. *Revista de Neuro-Psiquiatría*, 77(4), 262-270. <http://www.redalyc.org/articulo.oa?id=372033988009>
- Bonilla Sepúlveda, O. A. (2009). *Madres adolescentes: perfil epidemiológico y asociación con trastornos de depresión, ansiedad y adaptativo*. [Trabajo de

- grado de especialización, Universidad CES]. <https://repository.ces.edu.co/handle/10946/816>
- Bustos, V. & Russo, A. (2018). Salud mental como efecto del desarrollo psicoafectivo en la infancia. *Psicogente*, 21(39), 183-202. <http://doi.org/10.17081/psico.21.39.2830>
- Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection techniques. *Computers & Electrical Engineering*, 40(1), 16-28. <https://doi.org/10.1016/j.compeleceng.2013.11.024>
- Chikersal, P., Doryab, A., Tumminia, M., Villalba, D. K., Dutcher, J. M., Liu, X., Cohen, S., Creswell, K. G., Mankoff, J., Creswell, J. D., Goel, M., & Dey, A. K. (2021). Detecting depression and predicting its onset using longitudinal symptoms captured by passive sensing: A machine learning approach with robust feature selection. *ACM Transactions on Computer-Human Interaction*, 28(1), Article 3, 41 pages. <https://doi.org/10.1145/3422821>
- Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 297–334. <https://doi.org/10.1007/BF02310555>
- De Lacy, N., Ramshaw, M.J., McCauley, E. *et al.* (2023). Predicting individual cases of major adolescent psychiatric conditions with artificial intelligence. *Transl Psychiatry*, 13, 314. <https://doi.org/10.1038/s41398-023-02599-9>
- Dixit A. (2017). *Ensemble Machine Learning. A beginner's guide that combines powerful machine learning algorithms to build optimized models.* - Packt Publishing. -
- Durstewitz, D., Koppe, G. and Meyer-Lindenberg, A. (2019). Deep neural networks in psychiatry. *Mol Psychiatry*. 24, 1583-1598. <https://doi.org/10.1038/s41380-019-0365-9>
- Eaves RC., Darch C. & Williams TO Jr. (2004). Attention to novelty, fear-anxiety, and age: their effects on conduct problems. *The Journal of Genetic Psychology*, 165(4), 425–449. <https://doi.org/10.3200/GNTP.165.4.425-450>
- Erazo, O. & Martínez, J. (2024). Motivación hacia el aprendizaje, asociación con la depresión, la intimidación y sustancias psicoactivas en adolescentes escolarizados. *Revista Criminalidad*, 66(2), 167-182. <https://doi.org/10.47741/17943108.611>
- Fumero, A. & Navarrete, G. (2016). Personalidad y Malestar Psicológico: Aplicación de un Modelo de Redes Neuronales. *Revista Iberoamericana de Diagnóstico y Evaluación - e Avaliação Psicológica*, 1(41), 28-38. ISSN: 1135-3848. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=6474699>
- Gamboa, J & Salinas, J. (2022). Predicción de la situación académica en alumnos de pregrado usando algoritmos de machine learning. *Revista Perfiles*, 1(1), 04-10. ISSN: 2477-9105. DOI: http://scielo.senescyt.gob.ec/scielo.php?script=sci_arttext&pid=S2477-91052022000200004
- González-Hernández, J., Gómez-López, M., Pérez-Turpin, J. A., Muñoz-Villena, A., & Andreu-Cabrera, E. (2019). Perfectly active teenagers. When does physical exercise help psychological well-being in adolescents? *International Journal of Environmental Research and Public Health*, 16(22), 4525. <https://doi.org/10.3390/ijerph16224525>
- Güemes-Hidalgo, M., Ceñal, M.J. & Hidalgo Vicario, M. I. (2017). Desarrollo durante la adolescencia. Aspectos físicos, psicológicos y sociales. *Pediatría Integral*. XXI(4), 233-244. <https://www.pediatriaintegral.es/publicacion-2017-06/desarrollo-durante-la-adolescencia-aspectos-fisicos-psicologicos-y-sociales/>
- Hawes, M. T., Schwartz, H. A., Son, Y., & Klein, D. N. (2023). Predicting adolescent depression and anxiety from multi-wave longitudinal data using machine

- learning. *Psychological Medicine*, 53(13), 6205–6211. <https://doi.org/10.1017/S0033291722003452>
- Hervías Higuera, P., Gonzalvo Navarro, M.M., Moreno Fernández, A.M., Sánchez Mascaraque, P. & Correas Lauffer, J. (2022). Salud mental en la adolescencia (I). Ansiedad y depresión. *Medicine - Programa de Formación Médica Continuada Acreditado*, 13(61), 3581-3589. <https://doi.org/10.1016/j.med.2022.08.001>
- Instituto Nacional de la Salud Mental (2024). *Depresión*. Recuperado el 13 de noviembre de 2024 de https://www.nimh.nih.gov/sites/default/files/health/publications/espanol/depresion-sp/depresion_24.pdf
- James, G., Witten, D., & Hastie, T. (2013). *An introduction to statistical learning: With applications in R*. Springer. <https://doi.org/10.1007/978-1-0716-1418-1>
- Kashani JH, Nair SS, Rao VG, Nair J, Reid JC. (1996). Relationship of personality, environmental, and DICA variables to adolescent hopelessness: a neural network sensitivity approach. *Journal Am Acad Child Adolesc Psychiatry*. 35(5):640-645. <https://doi.org/10.1097/00004583-199605000-00019>
- Kilaskar, M., Saindane, N., Ansari, N. et al. (2022). Machine Learning Algorithms for Analysis and Prediction of Depression. *SN COMPUT. SCI*. 3, 103. <https://doi.org/10.1007/s42979-021-00967-0>
- Lantz B. (2019). *Machine Learning with R*. - Packt Pub.
- Leonangeli, S., Michelini, Y. & Rivarola Montejano, G. (2024). Depresión, Ansiedad y Estrés en Estudiantes Universitarios Antes y Durante Los Primeros Tres Meses de Cuarentena por COVID-19, *Revista Colombiana de Psiquiatría*, 53(3), 284-294. <https://doi.org/10.1016/j.rcp.2022.04.008>
- Lopez-Vega, J., Amaya-Gil, M. K., Salamanca Camargo, Y. & Caro Castillo, J. D. (2020). Relación entre psicopatologías e ideación suicida en adolescentes escolarizados de Colombia. *Psicogente*, 23(44), 1-18. <https://doi.org/10.17081/psico.23.44.3709>
- Martín, M.J Barba, C. Lázaro, J. Cuenca, E. Samaniego (2022). Prevención de los problemas socioemocionales en los centros educativos. Una propuesta de intervención psicoeducativa. *Revista de Orientación Educativa AOSMA*, 31, 36-66. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=8558466>
- Martínez de Velasco, S. R., Matas Ochoa, A., Banzo-Arguis, C., & Quintero, J. (2023). Protocolo diagnóstico y terapéutico de los trastornos de ansiedad. *Medicine - Programa de Formación Médica Continuada Acreditado*, 13(84), 4965-4969. <https://doi.org/10.1016/j.med.2023.08.005>
- Meylan, P., Schmidt, I. & Alam, M. (2024). Protegidos: Hacia un cambio de paradigma en la salud mental en Latinoamérica. *Revista Latinoamericana de Psiquiatría*, 23(1), 7-14. <https://www.webapal.org/docs/revista/revista-apal124.pdf>
- Mentis, A.F.A., Lee, D. and Roussos, P. (2023). Applications of artificial intelligence - machine learning for detection of stress: a critical overview. *Mol Psychiatry*. <https://doi.org/10.1038/s41380-023-02047-6>
- Ministerio de Salud y Protección Social (2015). *Encuesta Nacional de Salud Mental, Colombia 2015*. https://www.minjusticia.gov.co/programas-co/ODC/Publicaciones/Publicaciones/CO031102015-salud_mental_tomol.pdf
- Ministerio de Salud y Protección Social (2022). *Salud mental: asunto de todos*. Boletín de Prensa No. 481 de 2022. <https://www.minsalud.gov.co/Paginas/Salud-mental-asunto-de-todos.aspx>. [consultado el 15 de noviembre de 2024]
- Molina, M. F., Iribarne, K., Menéndez, M. A. & Álvarez Iturain A. R. (2023). La percepción de un futuro significativo y su relación con el autoconcepto y las experiencias positivas durante pandemia de covid-19 en estudiantes de nivel superior. *Psicogente*, 26(49), 1-24. <https://doi.org/10.17081/psico.26.49.5765>

- Mora Cancino, A.M. & Hernández Valencia, M. (2015). Embarazo en la adolescencia: cómo ocurre en la sociedad actual, *Perinatología y Reproducción Humana*, 29(2), 76-82. <https://doi.org/10.1016/j.rprh.2015.05.004>
- Neff, E. (2022). Trastornos depresivos, *EMC - Tratado de Medicina*, 26(3), 1-9. ISSN 1636-5410. [https://doi.org/10.1016/S1636-5410\(22\)46841-3](https://doi.org/10.1016/S1636-5410(22)46841-3)
- Nickson, D., Meyer, C., Walasek, L., & Toro, C. (2023). Prediction and diagnosis of depression using machine learning with electronic health records data: a systematic review. *BMC medical informatics and decision making*, 23(1), 271. <https://doi.org/10.1186/s12911-023-02341-x>
- Núñez, J. C., Gázquez, J. C., Pérez-Fuentes, M. C., Molero, M. M., Martos, Á., Barragán, A. B., & Simón, M. M. (2017). *Psicología y Educación para la Salud*. Scinfinder.
- Organización Mundial de la Salud. (2023). *Depresión*. <https://www.who.int/es/news-room/fact-sheets/detail/depression> [consultado el 14 de noviembre de 2024].
- Organización Mundial de la Salud. (2024). *La salud mental de los adolescentes*. <https://www.who.int/es/news-room/fact-sheets/detail/adolescent-mental-health> [consultado el 13 de noviembre de 2024]
- Organización Panamericana de la Salud (2018). *Salud Mental*. <https://www.paho.org/es/temas/salud-mental>
- Ortiz-Hernández, L., & Valencia-Valero, R. G.. (2015). Disparidades en salud mental asociadas a la orientación sexual en adolescentes mexicanos. *Cadernos De Saúde Pública*, 31(2), 417-430. <https://doi.org/10.1590/0102-311X00065314>
- Otzen, Tamara, & Manterola, Carlos. (2017). Técnicas de Muestreo sobre una Población a Estudio. *International Journal of Morphology*, 35(1), 227-232. <https://dx.doi.org/10.4067/S0717-95022017000100037>
- Pardo A., G., Sandoval D., A., & Umbarila Z., D. (2004). Adolescencia y depresión. *Revista Colombiana de Psicología*, (13), 13-28. Recuperado de <https://www.redalyc.org/articulo.oa?id=80401303>
- Pinto, S.J. & Parente, M. (2024). Comprehensive review of depression detection techniques based on machine learning approach. *Soft Comput* 28, 10701-10725. <https://doi.org/10.1007/s00500-024-09862-1>
- Pocco, K. (2022). Un sistema experto para el diagnóstico del trastorno depresivo basado en redes neuronales. *Revista Cubana de Informática Médica*; 14(2): e519. <http://scielo.sld.cu/pdf/rcim/v14n2/1684-1859-rcim-14-02-e519.pdf>
- Quiceno, J. M., Gómez, A., Herrera, S. F., Vélez, A. M., Vinaccia, S. & Bahamón, M. J. (2022). Riesgo suicida y estrategias de afrontamiento al estrés en población campesina colombiana. *Psicogente*, 25(48), 1-19. <https://doi.org/10.17081/psico.25.48.5509>
- R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>
- Rodríguez De Avila, U., León Valle, Z. & Ceballos Ospino, G.. (2021). Ideación suicida, ansiedad, capital social y calidad de sueño en colombianos durante el primer mes de aislamiento físico por COVID-19. *Psicogente*, 24(45), 1-18. <https://doi.org/10.17081/psico.24.45.4075>
- Rodríguez, H. E., Hernández, A., Sánchez, C. M. & Gutiérrez C. (2022). Asociación entre el nivel de actividad física y los síntomas de ansiedad y depresión en estudiantes de la licenciatura en fisioterapia de la facultad de medicina de la UNAM. *Revista Mexicana de Fisioterapia* . ISSN: 2683-2887, 2:5. <https://remefis.com.mx/n%C3%BAmero-5>

- Rodríguez MR, Rincón HG, Velasco MM, Hernández CA, Ramos ML, González JA. (2012). Prevalencia de trastornos mentales en una población que demanda servicios de salud en una IPS de baja complejidad y su asociación con la ideación suicida la y discapacidad percibida. *Rev. Fac. Nac. Salud Pública*, 30(2), 141-151. <http://www.scielo.org.co/pdf/rfnsp/v30n2/v30n2a03.pdf>
- Rossi, R. (2001). *Para superar el estrés: Consejos del psicólogo*. Barcelona: De Vecchi.
- Salazar, E., Guzmán, K., & Munévar, F. (2016). Relación entre tipos de colegio y niveles de ansiedad en una muestra de escolares bogotanos. *Pensando Psicología*, 12(19), 69-79. <https://doi.org/10.16925/pe.v12i19.1329>
- Santomauro, Damian F et al. (2021). Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *The Lancet*, 398, Issue 10312, 1700 – 1712. [https://doi.org/10.1016/S0140-6736\(21\)02143-7](https://doi.org/10.1016/S0140-6736(21)02143-7)
- Shatte, A. B. R., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: a scoping review of methods and applications. *Psychological Medicine*, 49(9), 1426–1448. <https://doi.org/10.1017/S0033291719000151>
- Ursul, A., Herrera Guerra, E., & Galvan Patrignani, G. (2022). Riesgo de suicidio en adolescentes escolarizados: Suicidal risk and strategies for coping with stress in Colombian rural population. *Psicogente*, 25(48), 1-25. <https://doi.org/10.17081/psico.25.48.5422>
- Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace
- Wani, M. A., ELAffendi, M. A., Shakil, K. A., Imran, A. S., & Abd El-Latif, A. A. (2023). Depression screening in humans with AI and deep learning techniques. *IEEE Transactions on Computational Social Systems*, 10(4), 2074-2089. <https://doi.org/10.1109/TCSS.2022.3200213>
- Zhang, Z. (2024). Early warning model of adolescent mental health based on big data and machine learning. *Soft Comput*, 28, 811–828. <https://doi.org/10.1007/s00500-023-09422-z>

Esta obra está bajo: Creative commons attribution 4.0 international license. El beneficiario de la licencia tiene el derecho de copiar, distribuir, exhibir y representar la obra y hacer obras derivadas siempre y cuando reconozca y cite la obra de la forma especificada por el autor o el licenciente.

