

**Análisis estadístico descriptivo y predictivo de los intentos de suicidio en Colombia y la identificación de las ideas suicidas en redes sociales mediante modelos de aprendizaje profundo.<sup>1</sup>**

**Descriptive and predictive statistical analysis of suicide attempts in Colombia and the identification of suicidal ideation in social networks using deep learning models.**

Víctor Alfonso Guzmán Brand. Corporación Unificada Nacional de Educación Superior. Correo: [victora.guzman@cun.edu.co](mailto:victora.guzman@cun.edu.co)

 <https://orcid.org/0000-0002-6051-3153>

Laura Esperanza Gelvez García. Corporación Unificada Nacional de Educación Superior. Correo: [Laura\\_gelvez@cun.edu.co](mailto:Laura_gelvez@cun.edu.co)

 <https://orcid.org/0000-0003-0164-2972>

## Resumen

**Objetivo:** analizar desde la estadística descriptiva y predictiva los datos de intentos de suicidio en Colombia. Adicionalmente, desarrollar el estudio de datos no estructurados provenientes de una red social, empleando un modelo de aprendizaje profundo. **Metodología:** se emplea el proceso CRISP-DM (Cross Industry Standard Process for Data Mining) siendo un método probado para guiar proyectos de minería de datos, proporcionando un enfoque estructurado y eficiente. **Resultados:** los datos de intentos de suicidio muestran una tendencia ascendente, con un total acumulado de 237,924 casos reportados. En 2023, el promedio mensual fue de 3,404 intentos, lo

<sup>1</sup> Fecha de envío: 8 de mayo

Fecha de aceptación: 3 de junio

Fecha de publicación: 30 de junio

Cómo citar este artículo (APA): Guzmán Brand, V. A., & Gelvez García, L. E. Análisis estadístico descriptivo y predictivo de los intentos de suicidio en Colombia y la identificación de las ideas suicidas en redes sociales mediante modelos de aprendizaje profundo. *Revista De Ciencias Y Artes*, 3(1), 19-35.

<https://doi.org/10.37211/2789.1216.v3.n1.116>



que equivale a aproximadamente 111 casos diarios. Además, el análisis de sentimientos realizado sobre comentarios en la red social reveló que el 87.35% de estos presentan una alta probabilidad de estar asociados con ideación suicida o con una intención negativa. **Discusiones:** Las ideas suicidas presentes en las redes sociales reflejan una gran carga emocional. Estudios han señalado que un número creciente de personas utiliza estos espacios para manifestar sus sentimientos relacionados con el suicidio. **Conclusiones:** los hallazgos refuerzan la necesidad de implementar políticas de salud pública dinámicas, enfocadas en la prevención y en el monitoreo de grupos vulnerables, tanto en el entorno físico como digital.

**Palabras clave:** estadística; intento de suicidio; idea suicida; red social; aprendizaje profundo

## Abstract/Resumo

**Objective:** to analyze descriptive and predictive statistics on suicide attempt data in Colombia. Additionally, develop the study of unstructured data from a social network, using a deep learning model. **Methodology:** the CRISP-DM process (Cross Industry Standard Process for Data Mining) is used as a proven method to guide data mining projects, providing a structured and efficient approach. **Results:** data on suicide attempts show an upward trend, with a cumulative total of 237,924 reported cases. In 2023, the monthly average was 3,404 attempts, which equates to approximately 111 cases per day. In addition, sentiment analysis performed on social network comments revealed that 87.35% of these present a high probability of being associated with suicidal ideation or negative intent. **Discussion:** Suicidal ideation present on social networks reflects a high emotional charge. Studies have indicated that an increasing number of people use these spaces to express their feelings related to suicide. **Conclusions:** The findings reinforce the need to implement dynamic public health policies focused on prevention and monitoring of vulnerable groups, both in the physical and digital environment.

**Keywords:** statistic; suicide attempt; suicidal ideation; social network; deep learning.

## 1. Introducción

Cada año, más de 726,000 personas fallecen por suicidio, consolidándolo como una de las principales crisis de salud pública a nivel mundial. Este problema impacta de manera significativa a los jóvenes, posicionándose como la tercera causa de muerte entre quienes tienen entre 15 y 29 años. No obstante, este fenómeno no se limita a los países de altos ingresos; en 2021, el 73% de los suicidios ocurrió en naciones de ingresos bajos y medianos, evidenciando la gravedad del problema en áreas con escasos recursos para la salud mental (Organización Mundial de la Salud, 2024).

En Colombia, la tasa de intentos de suicidio por cada 100 000 habitantes en 2022 alcanzó 72,4, superando las cifras de años anteriores y, en particular, el récord histórico de 2019, que se situaba en 62,1. Al analizar la incidencia por sexo, las mujeres han presentado consistentemente tasas más altas durante todo el período de vigilancia. Sin embargo, al desglosar por sexo y grupo etario, se observan diferencias notables: las tasas más elevadas se registran en personas de entre 10 y 24 años en ambos sexos, con una mayor prevalencia en mujeres (Instituto Nacional de Salud, 2023).

Por otro lado, el suicidio es una tragedia que no solo acorta una vida, sino que deja profundas secuelas emocionales en familias, comunidades y sociedades. Sus causas son diversas y abarcan una compleja interacción de factores sociales, culturales, biológicos, psicológicos y ambientales, lo que dificulta su prevención (Benavides et al., 2019). A partir de lo expuesto, resulta crucial

enfocarse en la prevención del suicidio, con especial atención al manejo de los pensamientos y conductas suicidas. Para lograrlo, es fundamental adoptar intervenciones respaldadas por evidencia científica, lo que hace imprescindible continuar y fortalecer la investigación en este campo (Villafaña & Cárdenas, 2022).

Asimismo, la conducta suicida comprende diversas manifestaciones que varían tanto en su gravedad como en su nivel de riesgo. En su etapa más temprana, la ideación suicida se refiere a pensamientos sobre el deseo o la intención de morir, a menudo acompañados de fantasías sobre la propia muerte, aunque no siempre desembocan en acciones concretas (Campillo-Serrano & Fajardo-Dolci, 2021; Cortés et al., 2021). A continuación, la planificación suicida implica la elaboración de un plan específico para cometer el acto, definiendo métodos, lugares y momentos, lo que conlleva un mayor riesgo de evolución hacia el intento (Rangel & Cárdenas, 2022). El intento de suicidio, es un acto autoinfligido con el propósito de morir, aunque no siempre termina en un desenlace fatal (Vázquez et al., 2023). Finalmente, el suicidio representa la culminación mortal de dicho proceso, mediante cualquier método, con clara evidencia de la intención de causar la propia muerte (Ministerio de Salud y Protección Social, 2018).

En el ámbito del análisis descriptivo, diversas investigaciones han abordado los intentos de suicidio en Colombia. Un ejemplo es el estudio de Vigoya-Rubiano et al. (2019), que realizó un análisis descriptivo de los datos de intentos de suicidio en la ciudad de Tunja entre 2012 y 2017. De manera similar, Quemba et al (2022) examinaron la dinámica epidemiológica de los intentos de suicidio en niños y adolescentes durante el periodo 2016-2020.

En cuanto al análisis predictivo, Cifuentes et al. (2024) desarrollaron un modelo de pronóstico de intentos de suicidio en una ciudad colombiana, utilizando técnicas avanzadas. Además, estudios como los de Barajas-Aranda et al., (2024); Guzmán-Brand & Gelvez-García (2024) emplearon algoritmos de machine learning para identificar patrones y variables relevantes en los intentos de suicidio, lo que contribuye al enfoque predictivo mediante el uso de tecnologías emergentes.

El objetivo de este artículo es llevar a cabo un análisis estadístico descriptivo y predictivo sobre los datos de intentos de suicidio en Colombia entre los años 2016 - 2023, utilizando la información registrada en el Sistema Nacional de Vigilancia en Salud Pública (SIVIGILA). Adicionalmente, se propone desarrollar el estudio de datos no estructurados provenientes de una red social, empleando un modelo de aprendizaje profundo, con el fin de obtener una visión más completa y detallada del fenómeno del suicidio y su expresión en entornos digitales.

## **2. Metodología**

Para el proceso de analítica de datos se emplea el método denominado CRISP-DM (Cross Industry Standard Process for Data Mining) siendo probado para guiar proyectos de minería de datos, el cual proporciona un enfoque estructurado y eficiente (Espinosa-Zúñiga, 2020). Como metodología, detalla las fases clave de un proyecto, describe las tareas necesarias en cada una de estas etapas y explica las interrelaciones entre ellas (Mejía et al., 2024). Además, al funcionar como un modelo de proceso, ofrece una visión integral del ciclo de vida de la minería de datos, facilitando tanto la planificación como la ejecución de las actividades de análisis, lo que asegura una mayor coherencia y éxito en los resultados obtenidos (IBM, 2021). En la Figura 1, se expone la estructura de la analítica de datos:

**Figura 1.** Etapas y actividades en el proceso CRISP-DM (Cross Industry Standard Process for Data Mining)



**Nota:** en la gráfica se detalla cada fase y tarea desarrollada en el proyecto. Elaboración propia.

## 2.1 Compresión del problema

En esta etapa, se lleva a cabo una revisión de la literatura y los antecedentes relacionados con el tema, lo que permite obtener una comprensión clara y precisa de las características de los datos sobre los intentos de suicidio y la ideación suicida manifestada en las interacciones en redes sociales. Este análisis documental no solo proporciona un contexto sólido sobre los estudios previos, sino que también ayuda a identificar patrones, tendencias y vacíos en la información existente, que pueden ser útiles para orientar la investigación. Asimismo, ofrece una base teórica que facilita la interpretación de los datos y la formulación de hipótesis fundamentadas para los análisis descriptivos y predictivos.

## 2.2 Selección de los datos

Para la selección de los datos sobre los intentos de suicidio se toman los datos abiertos del Sistema Nacional de Vigilancia en Salud Pública (SIVIGILA), su responsabilidad radica en llevar a cabo un proceso de observación y análisis objetivo, sistemático y continuo de los eventos en salud. Esta plataforma proporciona una orientación, planificación, ejecución, monitoreo y evaluación de las intervenciones en salud pública, garantizando así una toma de decisiones informada y efectiva en la gestión sanitaria (Ministerio de Salud y Protección Social, 2024).

Para el análisis de datos no estructurados, se seleccionó la red social Reddit debido a su capacidad de proporcionar acceso a la información a través de una interfaz de programación de aplicaciones (API). Esta herramienta está disponible únicamente para desarrolladores que se registren en el sistema de seguridad de Reddit y acepten los términos y condiciones de uso de los datos. Una vez cumplidos estos requisitos, los datos son descargados y posteriormente filtrados en el entorno de desarrollo Google Colaboratory, lo que facilita su procesamiento y análisis en tiempo real.

## 2.3 Comprensión de datos

En esta fase de comprensión de datos, se emplea el lenguaje de programación Python, utilizando diversas librerías especializadas para la minería de datos. Entre ellas, destacan pandas y numpy para la manipulación y análisis de datos, matplotlib y seaborn para la visualización gráfica, y scikit-learn para la implementación de algoritmos de aprendizaje automático. Además, para el procesamiento del lenguaje natural, se integran librerías como nltk y spacy, que permiten el análisis tanto simbólico como estadístico del texto, facilitando una comprensión más profunda de los datos no estructurados.

## 2.4 Preparación

Para la preparación de los datos en el análisis de sentimientos de las interacciones en la red social, es fundamental llevar a cabo varios pasos de limpieza y normalización del texto. Inicialmente, se deben eliminar caracteres especiales, signos de puntuación y emoticones, ya que estos no aportan valor significativo. Este proceso se realiza utilizando expresiones regulares que permiten identificar y suprimir dichos elementos no deseados. Posteriormente, se aplican técnicas de corrección ortográfica, codificación de etiquetas y normalización de palabras para asegurar que el texto esté en un formato coherente y de alta calidad, optimizando así los resultados del análisis de sentimientos.

## 2.5 Consideraciones éticas

Este proceso investigativo se lleva a cabo conforme a lo establecido en la normativa colombiana sobre el manejo de datos y la inteligencia artificial, garantizando el cumplimiento de las disposiciones legales. En particular, se sigue lo estipulado en el Decreto 1743 de 2016, el cual regula el proceso de anonimización de microdatos. Según el artículo 2.2.3.1.1, dicho proceso consiste en la transformación técnica de los datos individuales de las unidades de observación, de manera que no sea posible identificar a los sujetos o características particulares de la fuente de información, preservando las propiedades estadísticas en los resultados (Función Pública, 2023).

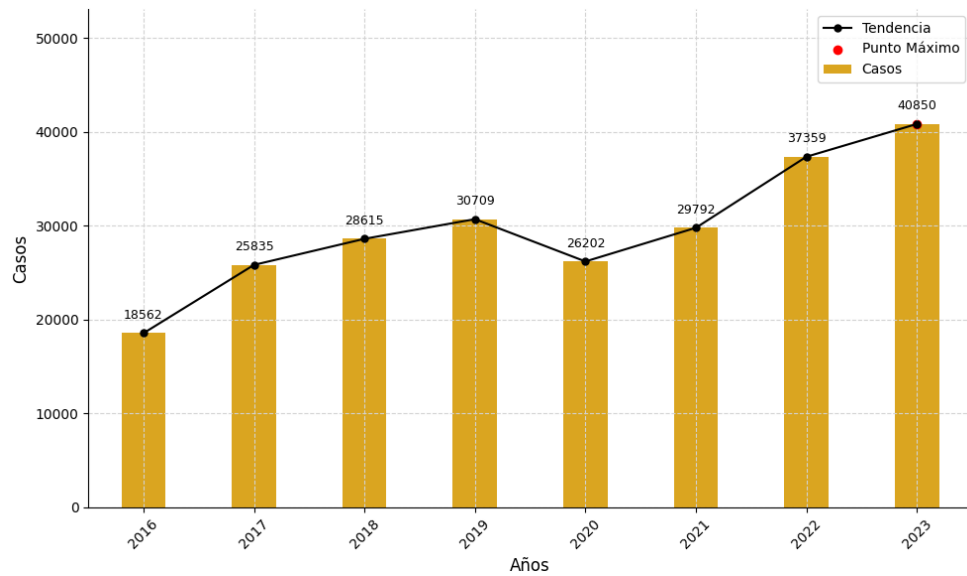
Además, se acata la Ley 1581 de 2012, conocida como la Ley de Protección de Datos Personales, que establece principios y obligaciones para el tratamiento de datos personales, incluyendo su recolección, uso y conservación (Secretaría Jurídica Distrital de la Alcaldía Mayor de Bogotá, 2024). En el contexto de la inteligencia artificial, se tiene en cuenta también el marco ético, que promueve su desarrollo y uso ético, fomentando la implementación de tecnologías que respeten los derechos fundamentales y aseguren la privacidad y la protección de los datos personales (Guio et al., 2021).

## 3. Resultados

### 3.1 Intentos de suicidio

El análisis de los intentos de suicidio registrados en el Sistema Nacional de Vigilancia en Salud Pública (SIVIGILA) en Colombia, entre 2016 - 2023, revela una tendencia creciente en los últimos años, con un total de 237,924 casos reportados durante este periodo. El año 2023 sobresale como el de mayor incidencia, representando el 17.1% del total de casos, seguido de 2022 con el 15.7%, 2019 con el 12.9%, 2021 con el 12.5%, 2018 con el 12.0%, 2020 con el 11.0%, 2017 con el 10.8% y 2016 con el 7.8%. Igualmente, se pasó de una tasa de intento de suicidio por cada 100,000 habitantes para el año 2016 de 39.64 a 78.09 para el año 2023. Además, en promedio durante 2023 se registraron 3,404 intentos de suicidio por mes, lo que equivale a 111 casos diarios. A continuación, se presentan los datos detallados en la Figura 2.

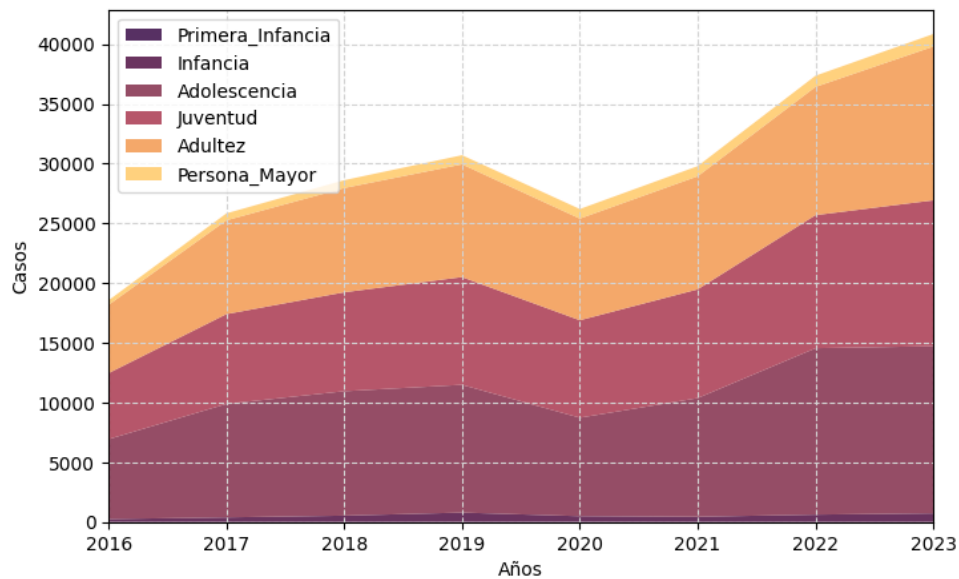
**Figura 2.** Comportamiento de los registros de intento de suicidio a lo largo de los años



**Nota:** en la gráfica se aprecia los casos de intento de suicidio a lo largo de los años mostrando una tendencia incremental. Elaboración propia.

En relación a la etapa del desarrollo con mayor afectación se tiene que es la adolescencia (12-17 años) con 83,390 casos, seguido de la adultez (27-59 años) con un 73,303, la juventud (18-26 años) con 70,947, persona mayor (60-más) con 6,031, infancia (6-11) 4,199 y primera infancia (0-5 años) con 54 casos. Sobre el género de la persona se tiene que el 63.7% es mujer y el 36.2% es hombre. Esto significa que por cada hombre que tiene un intento de suicidio, aproximadamente 1.7 mujeres también lo han intentado. En otras palabras, el número de mujeres con intentos de suicidio es un 76% mayor que el de los hombres. En la Figura 3 se muestran los datos mencionados:

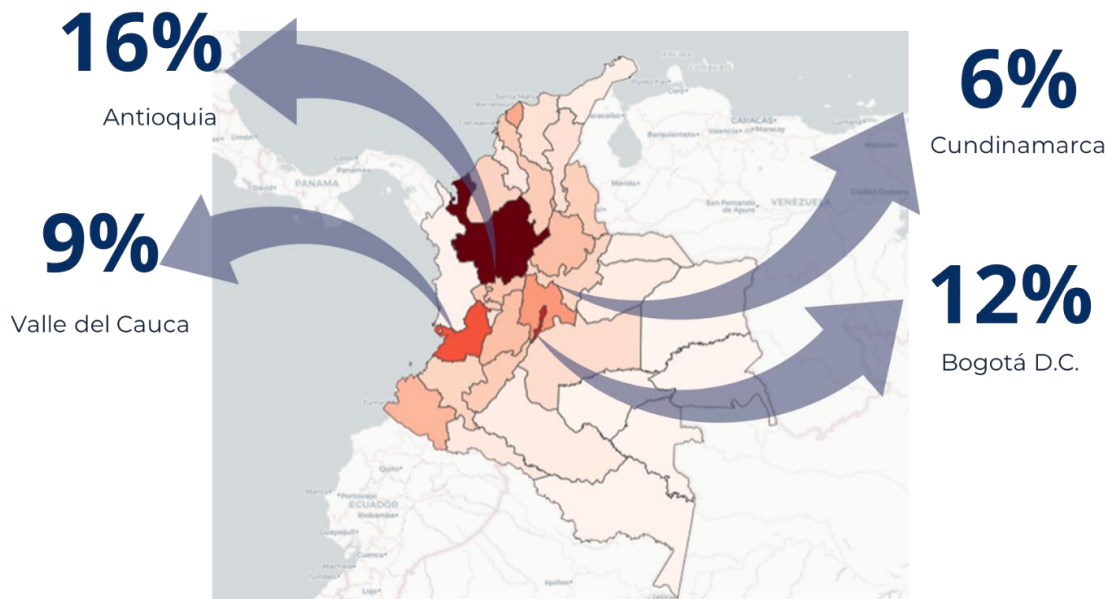
**Figura 3.** Etapas del desarrollo en razón a los casos de intento de suicidio



**Nota:** elaboración propia.

Afectando a grupos poblacionales como personas con trastornos psiquiátricos, privados de la libertad, personas migrantes y mujeres gestantes. Resto a la ubicación se tiene que los casos de intentos de suicidio se concentran en los departamentos de Antioquia con el 16%, Valle del Cauca 9%, Cundinamarca 6%, Atlántico 5% y Nariño 4%. Las ciudades con destacados números de casos esta Bogotá D.C. 12%, Medellín 7.4%, Cali 4.9%, Barranquilla 2.5% y Pasto 1.8%. Situación que se exhibe en la Figura 4.

**Figura 4.** Ubicación por departamentos de los casos de intento de suicidio



**Nota:** en el mapa de calor se observa los departamentos con mayor afectación. Elaboración propia.

### 3.2 Ideas suicidas expresadas en la red social

Para el análisis de las ideas suicidas expresadas en la red social Reddit, se recopilamos un total de 207,500 comentarios, los cuales fueron sometidos a un proceso de depuración para mejorar su claridad y facilitar su interpretación. Posteriormente, se llevó a cabo un estudio de sentimientos sobre estos datos, revelando que el 87.35% de los comentarios reflejan una alta probabilidad de estar vinculados con ideación suicida o una intencionalidad negativa. En contraste, el 12.65% restante sugiere una probabilidad elevada de expresar una intención positiva con baja presencia de ideación suicida. Esta clasificación mejora el rendimiento del modelo de aprendizaje profundo empleado en el estudio, aumentando su precisión para identificar patrones relevantes en el análisis de comentarios con contenido emocional crítico.

Simultáneamente, se realizó un análisis utilizando una nube de palabras sobre los comentarios, presentado en la Figura 5, lo que permitió identificar tendencias y patrones relevantes en la conversación del grupo social. Las palabras más frecuentes, como *suicidio* y *morir*, aparecieron de manera prominente y destacada en la visualización, lo que sugiere que estos términos son los ejes centrales de discusión entre los miembros del grupo. Este análisis cualitativo proporciona una visión clara de los temas predominantes que emergen en las interacciones, reforzando la importancia de estos conceptos en el diálogo del colectivo estudiado.



### ***B. Modelo BERT (Bidirectional Encoder Representations from Transformers)***

Este modelo de aprendizaje profundo se preentrena utilizando un vasto corpus de texto sin etiquetar que abarca toda la Wikipedia, equivalente a 2,500 millones de palabras. Este modelo se basa en redes neuronales Transformer, que emplean un mecanismo de atención para captar las relaciones contextuales entre palabras o subpalabras dentro de un texto. A diferencia de los modelos unidireccionales, BERT es profundamente bidireccional, lo que significa que, durante el entrenamiento, aprende información tanto del contexto anterior como del posterior a cada token. Esta bidireccionalidad es crucial para una comprensión más completa y precisa del significado del lenguaje, ya que permite al modelo considerar todo el contexto circundante al interpretar cada palabra (López-Condori & Gonzales-Saji, 2021).

Para el análisis predictivo utilizando el modelo BERT, a este se le aplica el proceso de transferencia de aprendizaje. Este enfoque permite que, tras su preentrenamiento en un amplio corpus de texto para comprender las reglas gramaticales y el contexto de las oraciones, ajuste su capacidad para identificar palabras y frases con intenciones o ideas suicidas dentro del texto. La transferencia de aprendizaje facilita que el modelo enfoque su atención en los elementos lingüísticos relevantes para detectar patrones asociados con la ideación suicida, aprovechando su entendimiento previo para mejorar la precisión en la clasificación y predicción de estos conceptos en nuevos datos.

Este proceso se enmarca en un conjunto o área de la inteligencia artificial denominada procesamiento del lenguaje natural, el cual busca replicar de manera artificial ciertos aspectos de la capacidad humana para el lenguaje, lo que abarca procesos de generación y comprensión del mismo (Báez et al., 2022). Este enfoque integra conocimientos provenientes de disciplinas como la lingüística, la psicología, la ingeniería y la informática. Además, se quiere desarrollar aplicaciones capaces de generar contenidos con un nivel de comprensión comparable al del lenguaje natural. Estas aplicaciones deben producir información bien estructurada y ser capaces de procesarla de manera eficiente (Moreira et al., 2021).

## **3.4. Evaluación de rendimiento**

### **A. Evaluación del modelo XGBoost**

La evaluación del rendimiento del modelo para la predicción de intentos de suicidio se lleva a cabo utilizando métricas específicas para series temporales. Estas métricas incluyen el error cuadrático medio (MSE), el error absoluto medio (MAE), la raíz del error cuadrático medio (RMSE) y el porcentaje de error medio absoluto (MAPE), las cuales permiten medir la precisión del modelo en términos de su capacidad para ajustar y prever los datos históricos. Además, se analiza el coeficiente de determinación ( $R^2$ ) para evaluar la proporción de variabilidad en los datos que el modelo puede explicar, proporcionando una visión integral de su eficacia predictiva. Datos que se exhiben en la Tabla 1.

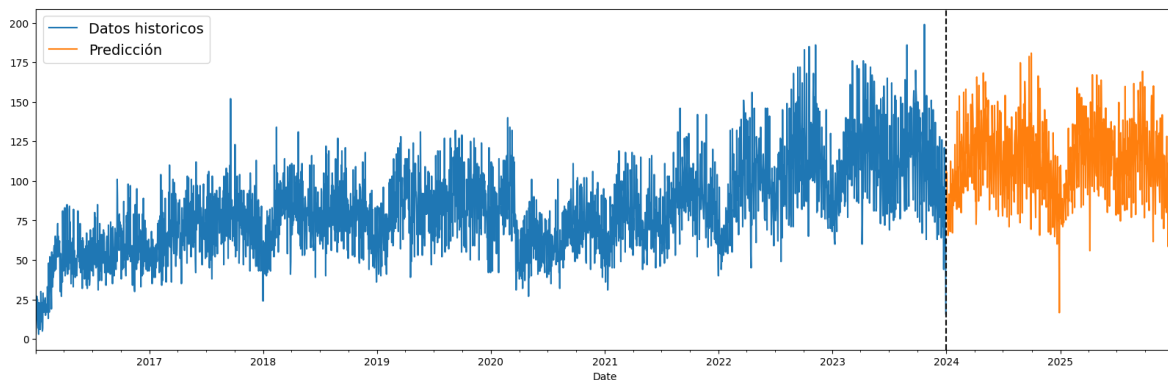
**Tabla 1.** Hiperparámetros del modelo XGBoost

Modelo	Hiperparámetros	MSE	MAE	RMSE	MAPE	R <sup>2</sup>
XGBoost	<ul style="list-style-type: none"> <li>• booster='gbtree'</li> <li>• seed=42</li> <li>• n_estimators=1000</li> <li>• early_stopping_rounds=10</li> </ul>	2.51	1.14	1.58	0.01	0.99
	<ul style="list-style-type: none"> <li>• objective='reg:squarederror'</li> <li>• reg_lambda=0.1</li> <li>• max_depth=6</li> <li>• eta=0.1</li> </ul>					

**Nota:** evaluación con las métricas para algoritmos de series temporales. Elaboración propia.

Tras el reentrenamiento con un total de 217,924 registros, el modelo evidencia un buen ajuste a las fluctuaciones de los datos, lo que permite proyectar los intentos de suicidio para los próximos dos años. Las predicciones indican un incremento en los casos proyectados para 2024, con un aumento del 0.71% (41,139) en comparación con los registros de 2023 (40,850). Para 2025, se anticipa un incremento adicional del 0.32% en relación con los datos históricos, subrayando una tendencia ascendente. Estudio que proveniente de la figura 7 así:

**Figura 7.** Análisis del comportamiento temporal de los casos y las predicciones generadas por el modelo.



**Nota:** comportamiento de los casos de intento de suicidio a lo largo del tiempo y la predicción por el algoritmo. Elaboración propia.

### B. Evaluación del modelo BERT

Para evaluar el rendimiento del modelo de aprendizaje profundo BERT, se comparó con otros algoritmos de machine learning, como Naive Bayes Kernel (NB) y Regresión Logística. Estos algoritmos son comúnmente utilizados en tareas de análisis de sentimientos, clasificación de documentos y correos electrónicos. Sin embargo, a diferencia de BERT, que captura el contexto y las sutilezas del lenguaje de manera bidireccional, estos métodos tradicionales no son capaces de interpretar las intenciones contextuales de las ideas presentes en el texto. Esto limita su capacidad para comprender de manera integral la complejidad y la profundidad del contenido textual. De tal manera, que los resultados son mostrados en la Tabla 2.

**Tabla 2.** Resultados de evaluación de los modelos

Algoritmo	Accuracy	F1 Score	ROC AUC	Recall
Naive Bayes Kernel - MultinomialNB	0.96	0.96	0.96	0.96
Regresión Logística	0.95	0.95	0.96	0.95
BERT (Bidirectional Encoder Representations from Transformers)	0.99	0.99	0.99	0.99

**Nota:** en la tabla se expone la evaluación y comparación de los algoritmos. Elaboración propia.

#### 4. Discusión

Los intentos de suicidio se han consolidado como un desafío alarmante para la salud pública, presentando una tendencia creciente a lo largo de los años y un repunte considerable en el periodo más reciente. Este incremento ha llamado la atención de las autoridades sanitarias, destacando la necesidad urgente de abordar el problema de manera integral. La evidencia indica que los factores psicosociales, en particular los problemas afectivos y las dificultades socioeconómicas, desempeñan un papel fundamental en la incidencia de los intentos de suicidio en la población analizada (Restrepo et al., 2023).

Igualmente, los resultados apuntan a que el segmento poblacional con mayor afectación es la adolescencia, seguido de la adultez y la juventud. Información respaldada por Quemba et al. (2022) el cual encuentra que los intentos de suicidio aumentan notablemente después de la pubertad. Las mujeres intentan suicidarse más que los hombres, pero los hombres tienen una mayor tasa de suicidio consumado (Castro-Moreno et al., 2023). La depresión mayor afecta a más del 50% de los adolescentes que se suicidan, y el alcoholismo es un factor importante en muchos casos de intentos de suicidio, asociándose a conductas violentas y accidentes fatales (Gómez-Restrepo et al., 2002).

De manera similar, el suicidio afecta significativamente a grupos vulnerables, como personas con trastornos psiquiátricos, población carcelaria, migrantes y mujeres gestantes. La prevalencia de intentos de suicidio es particularmente elevada en individuos con enfermedades mentales, especialmente depresión y ansiedad (Calderón-Ramírez et al., 2017; Pabon-Basto, 2021). La ideación suicida persistente es un fuerte indicador de futuros intentos de suicidio (Aparicio et al., 2020). Algunas comunidades presentando una mayor prevalencia de intentos de suicidio en comparación con otras (González, 2024).

Asimismo, afecta principalmente a los departamentos de Antioquia, Cundinamarca, Valle del Cauca y la ciudad de Bogotá. Además, vivir en zonas urbanas, especialmente en la capital, se ha identificado como un factor de riesgo importante (Murillo et al., 2022). Al igual, que Guzmán-Brand & Gélvez-García (2024) expone que la ubicación de los casos registrados, se observa que el 95% de los incidentes ocurrieron en la cabecera municipal, lo que indica una concentración significativa en las zonas urbanas. Además, los métodos varían según el género, en el caso de las mujeres utilizan más frecuentemente el envenenamiento y los hombres métodos más letales como armas de fuego y ahorcamiento (Beitia-Cardona et al., 2019).

El informe de la Defensoría del Pueblo (2024) destaca la insuficiencia de personal especializado para abordar situaciones relacionadas con la salud mental. Además, se señala que, en los departamentos de la amazonia, que registran las tasas elevadas de suicidio, no se cuenta con este tipo de profesionales. Además, en la infancia, diversos factores pueden aumentar el riesgo de

conductas suicidas, entre estos se incluyen el acoso escolar, la violencia intrafamiliar, la pobreza, la violencia sexual y de género, la discriminación, la exposición al conflicto armado, la exclusión social y la falta de acceso a servicios de salud mental adecuados (Save the Children, 2024).

Por otro lado, al aplicar el algoritmo XGBoost para proyectar la evolución de los casos en los próximos años, los resultados reflejan una tendencia al alza basada en el comportamiento histórico de los datos. Al respecto, los datos más recientes muestran una notable tendencia creciente en las tasas de suicidio (Ministerio de Salud y Protección Social, 2022). Acerca de las tasas de intentos de suicidio han mostrado una tendencia al alza desde finales de los años 90, con disparidades significativas según el género, la edad, el nivel educativo y la región. Con factores de riesgo como la ideación suicida persistente y los trastornos mentales (Cardona-Arango et al., 2016; Chaparro-Narváez et al., 2019).

En relación con las ideas suicidas indicadas en redes sociales, se observa que la mayoría de las conversaciones contienen expresiones de este tipo. Se ha señalado que cada vez más personas recurren a las redes sociales para manifestar sus sentimientos suicidas (Del Castillo et al., 2023). Además, los usuarios que comparten contenido relacionado con el suicidio muestran una alta conectividad recíproca, lo que sugiere la existencia de una comunidad virtual estrechamente vinculada (Durán-Mañes et al., 2023).

Por otra parte, el modelo de aprendizaje profundo denominado BERT muestra elevados puntajes en su evaluación de desempeño con un 0.99% de precisión, identificando de forma excepcional las ideas suicidas expresadas en el texto. Además, ha demostrado ser efectivo en la identificación de ideación suicida en datos de Twitter en portugués, alcanzando un 79% de puntuación F1 y menos del 9% de falsos negativos (De Carvalho et al., 2020). Igualmente, estos modelos superaron a otros algoritmos en la detección de intenciones suicidas en tweets, con una precisión del 95,39% en pruebas (Ananthakrishnan et al., 2022).

## 5. Conclusiones

El análisis de los intentos de suicidio registrados en el Sistema Nacional de Vigilancia en Salud Pública (SIVIGILA) entre 2016 y 2023 revela una preocupante tendencia creciente. El año 2023 sobresale por registrar el mayor número de casos, lo que sugiere un aumento significativo en la incidencia de este fenómeno en comparación con años anteriores. Los adolescentes, representan el grupo con mayor afectación, seguido de adultos y jóvenes. Esta distribución etaria refleja la vulnerabilidad de los adolescentes, una etapa crucial del desarrollo caracterizada por importantes desafíos emocionales y psicológicos. Además, las mujeres son el grupo demográfico más afectado, destacándose su exposición a factores de riesgo relacionados con el suicidio, lo que indica una problemática que podría estar asociada a condiciones socioeconómicas, psicológicas y culturales específicas.

Geográficamente, los departamentos de Antioquia, Valle del Cauca y Cundinamarca, así como las ciudades de Bogotá, Medellín y Cali, concentran la mayor cantidad de casos. Esta concentración en áreas urbanas de gran población sugiere que el fenómeno del suicidio puede estar relacionado con factores como el estrés urbano, la desigualdad social y el acceso limitado a servicios de salud mental.

El análisis de comentarios en redes sociales, específicamente en Reddit, ha permitido identificar un alto nivel de ideación suicida. A través de técnicas de procesamiento del lenguaje natural, se pudo extraer una serie de patrones recurrentes en los comentarios, entre los cuales destacan palabras clave como "suicidio" y "morir", lo que revela que las conversaciones en estas plataformas giran en torno a temáticas de gran carga emocional. Esta información es de vital

importancia para entender cómo los individuos expresan su angustia y cómo las redes sociales pueden convertirse en un espacio tanto de vulnerabilidad como de posible intervención.

En el campo del análisis predictivo, el uso del modelo XGBoost ha permitido proyectar la evolución de los intentos de suicidio para los próximos dos años, lo que señala un incremento en los casos estimados. La estabilidad y confiabilidad del modelo fueron validadas mediante la prueba Dickey-Fuller, que confirma que la serie temporal es adecuada para el análisis y proyección. Este tipo de modelado predictivo ofrece una herramienta valiosa para las autoridades de salud pública, quienes podrían utilizar estos resultados para implementar estrategias preventivas basadas en tendencias futuras.

Por otro lado, el modelo BERT, basado en aprendizaje profundo, ha demostrado ser eficiente para la detección de ideación suicida en textos provenientes de redes sociales. A diferencia de otros algoritmos, BERT capta con precisión el contexto y la intención detrás de las palabras, lo que le permite identificar patrones lingüísticos asociados a pensamientos suicidas con un nivel de detalle y exactitud superior. Este modelo, al ser aplicado en un entorno de análisis de textos no estructurados, abre la puerta a nuevas formas de monitoreo de la salud mental en plataformas digitales, proporcionando un enfoque proactivo para la identificación temprana de comportamientos de riesgo.

En consecuencia, el presente estudio pone en evidencia la creciente incidencia de los intentos de suicidio en Colombia y ofrece una perspectiva integral para su análisis, combinando métodos para la minería de datos. Los hallazgos refuerzan la necesidad de implementar políticas de salud pública más dinámicas, enfocadas en la prevención y en el monitoreo de grupos vulnerables, tanto en el entorno físico como digital. Asimismo, se destaca el potencial de los algoritmos de aprendizaje automático como herramientas para anticipar comportamientos de riesgo y mejorar la intervención en salud mental.

## CONTRIBUCIÓN DE LOS AUTORES

A continuación, se describen los aportes de cada uno de los autores que contribuyeron al desarrollo de la investigación, según cada una de las categorías de la taxonomía CREDIT (<https://credit.niso.org/>)

Contribución	Apellido, Nombre de nombre de autor o autores que ha contribuido en cada categoría
<b>Escritura:</b> visualización, revisión y edición de las diferentes versiones, primer borrador.	Laura Gelvez García
<b>Gestión:</b> validación, supervisión, administración o planificación de la investigación	Laura Gelvez García
<b>Recursos y tecnología:</b> Fondos, recursos y materiales, software.	Víctor Guzmán Brand
<b>Investigación:</b> Análisis, conducción, curación de datos, metodología, conceptualización.	Víctor Guzmán Brand

### **DECLARACIÓN DE DISPONIBILIDAD DE DATOS**

Los autores declaran que los datos utilizados en la investigación están disponibles en su totalidad en el manuscrito.

### **DECLARACIÓN DE CONFLICTO DE INTERÉS**

Los autores declaran haber respetado los principios éticos de investigación y estar libre de cualquier conflicto de intereses en cuanto a la realización de la investigación y la divulgación de sus resultados.

### **DECLARACIÓN SOBRE EL USO DE INTELIGENCIA ARTIFICIAL**

Los autores declaran que el uso de la inteligencia artificial generativa y no generativa se ajusta a lo establecido en las normas para autores.

### **Referencias**

- Ananthkrishnan, G., Jayaraman, A., Trueman, T., Mitra, S., K, A., & Murugappan, A. (2022). Suicidal Intention Detection in Tweets Using BERT-Based Transformers. *2022 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, 1(1), 322-327. 2022 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS). <https://doi.org/10.1109/ICCCIS56430.2022.10037677>
- Aparicio, Y., Blandón, A. M., & Chaves, N. (2020). Alta prevalencia de dos o más intentos de suicidio asociados con ideación suicida y enfermedad mental en Colombia en 2016. *Revista Colombiana de Psiquiatría*, 49(2), 96-101. <https://doi.org/10.1016/j.rcp.2018.07.001>
- Báez, P., Paz, A., Chaparro, M., & Bucarey, T. (2022). Procesamiento de lenguaje natural para texto clínico en español: El caso de las listas de espera en Chile. *Revista Médica Clínica Las Condes*, 33(6), 576-582. <https://doi.org/10.1016/j.rmclc.2022.10.002>
- Barajas-Aranda, D., Torres-Soto, A., Torres-Soto, M., & Sicilia-Urbán, M. (2024). Comparison and explainability of machine learning models in predictive suicide analysis. *DYNA New Technologies Journal*, 11(1), 1-25. <https://revista-dyna.com/dyna-newtech/index.php/dyna-newtech/article/view/35>
- Beitia-Cardona, P. N., Rodríguez-Torres, E., Estrada-González, C., & Benavides-Portilla, M. (2019). Intento suicida y factores asociados en dos instituciones de Cali – Colombia. *Revista Ciencia y Cuidado*, 16(1), 19-31. <https://doi.org/10.22463/17949831.1542>
- Benavides, V., Villota, N., & Villalobos, F. (2019). Conducta suicida en Colombia: Una revisión sistemática. *Revista de Psicopatología y Psicología Clínica*, 24(3), Article 3. <https://doi.org/10.5944/rppc.24251>
- Calderón-Ramírez, V., Alcocer-Olaciregui, A., & Vargas-Moranth, R. (2017). Intentos de suicidio por intoxicación con sustancias químicas en Colombia. 2007—2013. *Duazary*, 14(2), 20-45. <https://doi.org/10.21676/2389783X.1963>
- Campillo-Serrano, C., & Fajardo-Dolci, G. (2021). Prevención del suicidio y la conducta suicida. *Gaceta médica de México*, 157(5), 564-569. <https://doi.org/10.24875/gmm.21000205>

- Cardona-Arango, D., Medina-Pérez, Ó. A., & Cardona-Duque, D. (2016). Caracterización del suicidio en Colombia, 2000-2010. *Revista Colombiana de Psiquiatría*, 45(3), 170-177. <https://doi.org/10.1016/j.rcp.2015.10.002>
- Castro-Moreno, L., Fuertes-Valencia, L., Pacheco-García, O., & Muñoz-Lozada, C. (2023). Factores De Riesgo Relacionados Con Intento de Suicidio Como Predictores De suicidio, Colombia 2016-2017☆. *Revista Colombiana de Psiquiatría*, 52(3), 176-184. <https://doi.org/10.1016/j.rcp.2021.03.002>
- Chaparro-Narváez, P., Díaz-Jiménez, D., & Castañeda-Orjuela, C. (2019). Tendencia de la mortalidad por suicidio en las áreas urbanas y rurales de Colombia, 1979-2014. *Biomédica*, 39(2), 20-35. <https://doi.org/10.7705/biomedica.v39i3.4427>
- Cifuentes, A., Méndez, T., & Jiménez, M. (2024). Análítica predictiva como apoyo en la salud pública: Modelos de pronóstico en salud mental con series de tiempo. *Revista EIA*, 21(42), 4227 pp. 1-22. <https://doi.org/10.24050/reia.v21i42.1763>
- Cortés, A., Román, M., Suárez, R., & Alonso, R. M. (2021). Conducta suicida, adolescencia y riesgo. *Anales de la Academia de Ciencias de Cuba*, 11(2), 1-16.
- De Carvalho, V., Giacón, B., Nascimento, C., & Nogueira, B. (2020). *Machine Learning for Suicidal Ideation Identification on Twitter for the Portuguese Language* (R. Cerri & R. Prati, Eds.; Vol. 12319, pp. 536-550). Springer International Publishing. [https://doi.org/10.1007/978-3-030-61377-8\\_37](https://doi.org/10.1007/978-3-030-61377-8_37)
- Defensoría del Pueblo. (2024). *Defensoría del Pueblo pide priorizar la salud mental de niñas, niños y adolescentes*. Defensoría. <https://www.defensoria.gov.co/-/defensor%C3%ADa-del-pueblo-pide-priorizar-la-salud-mental-de-ni%C3%B1as-y-adolescentes>
- Del Castillo, R., García-Murillo, L., Mallol-Castaño, L., Calderón, P., Aranda, A., & Palanca-Maresca, I. (2023). Redes sociales y conductas suicidas en la infancia y la adolescencia durante la pandemia de COVID-19: Una relación difícil de estimar. *Revista de Psiquiatría Infanto-Juvenil*, 40(3), 0-45. <https://doi.org/10.31766/revpsij.v40n3a2>
- Durán-Mañes, A., Paniagua-Rojano, F., & Fernández-Beltrán, F. (2023). Análisis de medios y audiencias en redes sociales ante informaciones sobre suicidio. *Revista Comunicar*, 2(2), 1-20. <https://www.revistacomunicar.com/html/77/es/77-2023-10.html>
- Espinosa-Zúñiga, J. (2020). Aplicación de metodología CRISP-DM para segmentación geográfica de una base de datos pública. *Ingeniería, investigación y tecnología*, 21(1), 1-20. <https://doi.org/10.22201/fi.25940732e.2020.21n1.008>
- Función Pública. (2023). *Decreto 1743 de 2016*. <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=77734>
- Gómez-Restrepo, C., Malagón, N., Bohórquez, A., Díaz-Granados, F., García, M., & Fernández, C. (2002). Factores asociados al intento de suicidio en la población colombiana. *Revista Colombiana de Psiquiatría*, 31(30), 283-298. [http://www.scielo.org.co/scielo.php?script=sci\\_arttext&pid=S0034-74502002000400002](http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0034-74502002000400002)
- González, C. (2024). Factores diferenciados según sexo del intento suicida en infantes menores de 11 años en Boyacá, Colombia, entre los años 2018 y 2022. *Revista Colombiana de Psiquiatría*, 2(2), 1-20. <https://doi.org/10.1016/j.rcp.2024.03.009>

- Guio, A., Tamayo, E., & Gómez, P. (2021). *Marco ético para la inteligencia artificial en Colombia* (No. Versión 1; pp. 1-63). Departamento Administrativo de la Presidencia de la República. <https://minciencias.gov.co/sites/default/files/marco-etico-ia-colombia-2021.pdf>
- Guzman-Brand, V., & Gelvez-García, L. (2024). Análisis Estadístico y Predictivo de los Datos de Eventos, Víctimas y Desminado Humanitario de las Minas Antipersonal (MAP) en Colombia. *Ciudad Paz-ando*, 17(1), 100-120. <https://doi.org/10.14483/2422278X.21706>
- Guzmán-Brand, V., & Gélvez-García, L. (2024). Identificación de patrones a través de algoritmos de machine learning en los casos registrados de intentos suicidas en una ciudad de Colombia. *Psicoespacios*, 18(32), 50-65. <https://doi.org/10.25057/21452776.1634>
- IBM. (2021, agosto 17). *Conceptos básicos de ayuda de CRISP-DM*. <https://www.ibm.com/docs/es/spss-modeler/saas?topic=dm-crisp-help-overview>
- Instituto Nacional de Salud. (2023). *Informe de Evento Intento de Suicidio, 2022* (pp. 1-21). Instituto Nacional de Salud. <https://www.ins.gov.co/buscador-eventos/Informesdeevento/INTENTO%20DE%20SUICIDIO%20INFORME%202022.pdf>
- López-Condori, J., & Gonzales-Saji, F. (2021). Análisis de sentimiento de comentarios en español en Google Play Store usando BERT. *Ingeniare. Revista chilena de ingeniería*, 29(3), 557-563. <https://doi.org/10.4067/S0718-33052021000300557>
- Mejía, C., Rincón, M., Palmera, L., & Arevalo, L. (2024). Aplicación de machine learning y metodología CRISP-DM para la clasificación precisa de severidad en casos de dengue. *Revista Colombiana De Tecnologías De Avanzada (RCTA)*, 1(43), 45-65. <https://doi.org/10.24054/rcta.v1i43.2822>
- Ministerio de Salud y Protección Social. (2018). *Boletín de salud mental Conducta suicida Subdirección de Enfermedades No Transmisibles* (pp. 1-25). MINSALUD. <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/ENT/bolet-in-conducta-suicida.pdf>
- Ministerio de Salud y Protección Social. (2022). *Reportes Salud Mental*. <https://rssvr2.sispro.gov.co/ObsSaludMental/>
- Ministerio de Salud y Protección Social. (2024). *Sistema de Vigilancia en Salud Pública* [Publica]. Ministerio de Salud y Protección Social. <https://www.minsalud.gov.co/salud/Paginas/SIVIGILA.aspx>
- Moreira, D., Cruz, I., González, K., Quirumbay, A., Magallan, C., Guarda, T., Andrade, A., & Castillo, C. (2021). Análisis del estado actual de procesamiento de lenguaje natural. *RISTI: Revista Ibérica de Sistemas e Tecnologías de Informação*, 42(42), 126-136. <https://dialnet.unirioja.es/servlet/articulo?codigo=8624557>
- Murillo, L., Quemba, M., Vargas, L., Florez, I., & Contreras, J. (2022). Epidemiological behavior of suicide attempt in Colombian adolescents years 2016-2019: An ecological study. *Revista Latino-Americana de Enfermagem*, 30(15), 1-25. <https://doi.org/10.1590/1518-8345.6240.3807>
- Organización Mundial de la Salud. (2024). *Suicidio*. <https://www.who.int/es/news-room/fact-sheets/detail/suicide>

- Pabón-Basto, A. (2021). Intentos de suicidio y Trastornos mentales. *Revista Habanera de Ciencias Médicas*, 20(4), 40-55. <https://revhabanera.sld.cu/index.php/rhab/article/view/3967>
- Pérez-Martínez, L., Tápanes-Suárez, E., Santos-Pérez, O., Cabrera-Hernández, J., & Nogueira-Rivera, D. (2021). Procedimiento para Índice Sintético de Gestión Ambiental: Validación con minería de datos. *Ingeniería Industrial*, 42(2), 1-20. [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1815-59362021000200060](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1815-59362021000200060)
- Quemba, M., Herrera, J., Mendoza, A., & Mendoza, B. (2022). Comportamiento epidemiológico del intento de suicidio en niños y adolescentes, Colombia 2016-2020. *Pediatría*, 55(1), 120-135. <https://doi.org/10.14295/rp.v55i1.355>
- Rangel, J., & Cárdenas, S. (2022). Definición de suicidio y de los pensamientos y conductas relacionadas con el mismo: Una revisión. *Psicología y Salud*, 32(1), 50-65. <https://doi.org/10.25009/pys.v32i1.2709>
- Restrepo, M., Mejía, S., & García, J. (2023). El intento suicida en relación con la edad y otras variables psicosociales. Un análisis de contexto. *Revista Virtual Universidad Católica del Norte*, 70(69), Article 69. <https://doi.org/10.35575/rvucn.n69a9>
- Save the Children. (2024). *La Prevención del Suicidio en la Niñez: Un Llamado a la Acción*. <https://savethechildren.org.co/la-prevencion-del-suicidio-en-la-ninez-un-llamado-a-la-accion/>
- Secretaría Jurídica Distrital de la Alcaldía Mayor de Bogotá. (2024). *Ley 1581 de 2012 Congreso de la República de Colombia*. Secretaría Jurídica Distrital de la Alcaldía Mayor de Bogotá. <https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=49981>
- Vázquez, P., Armero, P., Martínez, L., García, J. M., Bonet, C., Notario, F., Sánchez, A. R., Rodríguez, P. J., & Díez, A. (2023). Autolesiones y conducta suicida en niños y adolescentes. Lo que la pandemia nos ha desvelado. *Anales de Pediatría*, 98(3), 204-212. <https://doi.org/10.1016/j.anpedi.2022.11.006>
- Vigoya-Rubiano, A., Osorio-Barajas, Y., & Salamanca-Camargo, Y. (2019). Caracterización del intento de suicidio en una ciudad colombiana (2012 - 2017). *Duazary*, 16(1), 53-66. <https://www.redalyc.org/journal/5121/512162369007/html/>
- Villafaña, J., & Cárdenas, S. (2022). Definición de suicidio y de los pensamientos y conductas relacionadas con el mismo: Una revisión. *Psicología y Salud*, 32(1), 1-20. <https://doi.org/10.25009/pys.v32i1.2709>