
Identificación de patrones a través de algoritmos de machine learning en los casos registrados de intentos suicidas en una ciudad de Colombia

psycospacios

Identification of patterns through machine learning algorithms in registered cases of suicide attempts in a Colombian city

Guzmán Brand, Víctor Alfonso; Gélvez García, Laura Esperanza

 Víctor Alfonso Guzmán Brand

victora.guzman@cun.edu.co

Corporación Unificada Nacional de Educación Superior, Colombia

 Laura Esperanza Gélvez García

laura_gelvez@cun.edu.co

Corporación Unificada Nacional de Educación Superior, Colombia

Psicoespacios

Institución Universitaria de Envigado, Colombia

ISSN: 2145-2776

Periodicidad: Frecuencia continua

vol. 18, núm. 32, 1-15, 2024

psicoespacios@iue.edu.co

Recepción: 22 Febrero 2024

Aprobación: 03 Mayo 2024

Publicación: 16 Mayo 2024

URL: <http://portal.amelica.org/amelijournal/743/7434694005/>

DOI: <https://doi.org/10.25057/21452776.1634>

Resumen: Objetivo: identificar patrones en los casos de intentos suicidas reportados en el Sistema Nacional de Vigilancia en Salud Pública por la alcaldía de Bucaramanga, Colombia, durante el período 2016-2023. Metodología: se emplea un enfoque cuantitativo con el propósito de describir, explicar y predecir los fenómenos estudiados, utilizando el método Knowledge Discovery in Databases para guiar el proceso de minería de datos. Resultados: se muestra una alta concentración de intentos suicidas, especialmente en el año 2022, por lo que se subraya la importancia de implementar estrategias específicas para diferentes grupos de edad y género. Discusión: se destaca que los jóvenes representan el 40,5 % de los casos, con un riesgo aumentado en el trascurso de la pandemia, especialmente entre aquellos con trastornos mentales. Las mujeres constituyen el 60,5 % de los casos, mayormente en áreas urbanas y en el régimen contributivo de salud. Conclusiones: se enfatiza en la importancia de considerar los antecedentes de intentos previos, las tensiones con la pareja y el uso de sustancias psicoactivas como factores de riesgo. Además, se resalta la conexión entre trastornos mentales e intentos suicidas, señalando la intoxicación como el método más común. Se subraya un riesgo mayor en personas vulnerables a diversas formas de violencia y condiciones socioeconómicas precarias. El método Random Forest demuestra ser efectivo para identificar factores clave, mientras que los algoritmos aplicados de *machine learning* podrían mejorar la detección temprana.

Palabras clave: patrones, algoritmo, *machine learning*, intento suicida, población.

Abstract: Objective: to identify patterns in the cases of suicide attempts reported in the National Public Health Surveillance System by the mayor's office of the city of Bucaramanga, Colombia, during the period 2016-2023. Methodology: a quantitative approach is employed with the purpose of describing, explaining and predicting the phenomena studied, using the Knowledge Discovery in Databases method to guide the data mining process. Results: a high concentration of suicide attempts is shown, especially in the year 2022, underlining the importance of specific strategies for different age and gender groups. Discussion: it is highlighted that young people represent 40.5% of the cases, with an increased risk in the

course of the pandemic, especially among those with mental disorders. Women account for 60.5% of cases, mostly in urban areas and in the contributory health care system. Conclusions: the importance of considering a history of previous attempts, tensions with the partner and the use of psychoactive substances as risk factors is emphasized. In addition, the connection between mental disorders and suicide attempts is highlighted, pointing out intoxication as the most common method. A higher risk is highlighted in people vulnerable to various forms of violence and precarious socioeconomic conditions. The Random Forest method proves to be effective in identifying key factors, while applied machine learning algorithms could improve early detection.

Keywords: patterns, algorithm, machine learning, suicide attempt, population.

INTRODUCCIÓN

El comportamiento suicida constituye un proceso dinámico que abarca aspectos emocionales, cognitivos y conductuales. Esta problemática, de naturaleza multifactorial, sigue un curso que se puede definir con distintas fases y categorías de análisis (Al-Halabí & Fonseca-Pedrero, 2023). El proceso suicida se suele desarrollar en una secuencia progresiva que incluye pensamientos, planes y múltiples intentos, con un aumento gradual en la letalidad, hasta llegar al suicidio consumado (Ministerio de Salud y Protección Social de Colombia, 2018).

Según la Organización Mundial de la Salud (2021), alrededor de 703 000 personas mueren por suicidio cada año, y un número aún mayor intenta suicidarse. Cada uno de estos actos representa una tragedia que afecta profundamente a familias, comunidades y naciones, dejando efectos duraderos en quienes están cerca de la víctima. El suicidio no hace distinciones de edad y en 2019 se situó como la cuarta causa de muerte más importante entre las personas de 15 a 29 años en el mundo.

Igualmente, son evidentes los factores que se vinculan a la conducta suicida, entre los que se destacan el maltrato infantil, los problemas académicos o conductuales en el contexto escolar, la depresión (Papalia & Martonell, 2017), el trastorno bipolar, la esquizofrenia (Restrepo Tobón et al., 2023), la ansiedad (Sindeev et al., 2019), el ciber acoso, los problemas familiares, la soledad, las dificultades económicas (Myers & Twenge, 2019). El sexo también es un influyente, pues las mujeres presentan mayores porcentajes de ideación e intentos suicidas (Morfin López & Ibarra López, 2018); además, otros factores son la planificación del hecho (Mojica Walteros et al., 2023), el estrés crónico y el haber presentado intentos previos (Castro-Osorio et al., 2022; Paredes Lascano et al., 2022).

Por otro lado, los factores de riesgo más relevantes para el suicidio en adolescentes son: transición adolescente, cambios en el estilo de vida, influencia de grupos sociales y consumo de sustancias (Soto Salcedo et al., 2020). Otros factores son la baja tolerancia a la frustración, las limitadas habilidades de manejo de estrés, el control de impulsos deficiente, los problemas familiares, los eventos estresantes como desamores o abusos (Vigoya-Rubiano et al., 2019) y el uso inadecuado de tecnología (Robles Sánchez et al., 2020).

Los intentos de suicidio son más comunes entre mujeres, especialmente durante la adolescencia y la juventud. Los métodos utilizados incluyen la intoxicación, especialmente con plaguicidas en países en desarrollo, así como el uso de medicamentos, seguido del corte (Hinojosa-Vega et al., 2023). Las repeticiones de intentos pueden ocurrir dentro de los tres a seis meses siguientes al primer intento, o incluso durante el mismo año calendario. La historia previa de intentos suicidas es un factor de riesgo significativo para el suicidio consumado en la población en general (Mosquera Bahamón, 2020).

La aplicación de inteligencia artificial para detectar patrones en los intentos suicidas ha demostrado ser prometedora. Al emplear el algoritmo Random Forest, se pueden identificar variables relevantes que ayudan a orientar los factores de riesgo. Además, los algoritmos de clasificación aplicados a los datos muestran un buen ajuste, lo que facilita la automatización de procesos y la identificación de características importantes para la toma de decisiones. Del mismo modo, según Navarro Acosta et al. (2021), en el ámbito de la psicometría se ha evidenciado que el uso de técnicas de aprendizaje automático conlleva una mejora significativa en la descripción de instrumentos psicológicos. En particular, se destaca el comportamiento excepcional del algoritmo de máquinas de soporte vectorial en este contexto.

También, se ha incursionado en el estudio de la conducta suicida mediante el análisis de las interacciones en redes sociales, enfocándose en la detección de patrones en las ideas expresadas en textos, los cuales son analizados en términos de sentimientos utilizando técnicas de procesamiento del lenguaje natural (Castillo-Zúñiga et al., 2022). Además, se ha llevado este enfoque un paso más allá al emplear *chatbots* como asistentes para la preevaluación de pacientes (Biscaia Fernández et al., 2023).

Por lo tanto, se realiza un análisis exploratorio y descriptivo de la información, destacando el papel crucial de las técnicas de visualización de datos, las cuales permiten a los investigadores identificar patrones, anomalías y tendencias de manera más efectiva a través de representaciones visuales. Este enfoque simplifica la comprensión de la información contenida en los datos (Avila Pérez, 2021). Posteriormente, esta fase evoluciona hacia el proceso de modelado, en el que se aplica la inteligencia artificial con el fin de descubrir patrones previamente desconocidos pero válidos, novedosos, potencialmente útiles y comprensibles (Joyanes Aguilar, 2019).

El presente estudio tiene como objetivo identificar patrones mediante algoritmos de *machine learning* en los casos de intentos suicidas reportados en el Sistema Nacional de Vigilancia en Salud Pública (sivigila) por la alcaldía de la ciudad de Bucaramanga, en Colombia, entre los años 2016-2023, información anonimizada y registrada en el portal de datos abiertos.

METODOLOGÍA

Se toma la aproximación cuantitativa, la cual resulta idónea para estimar magnitudes o incidencia de fenómenos, así como para probar hipótesis (Baena Paz, 2014). Los datos se presentan en forma numérica y su recolección se basa en la medición, usualmente de variables relacionadas con las hipótesis planteadas. Esta recolección se lleva a cabo utilizando métodos estandarizados y aceptados por la comunidad científica (Hernández Sampieri et al., 2019). El objetivo último de los estudios cuantitativos es describir, explicar y predecir los fenómenos investigados, buscando identificar regularidades y relaciones causales entre las variables (Domínguez Granda, 2015). Si se sigue meticulosamente el proceso y se aplican reglas lógicas, los datos generados cumplirán con los estándares de validez y confiabilidad deseados, y las conclusiones derivadas contribuirán al avance del conocimiento científico (Caminotti & Toppi, 2020).

ANÁLISIS DE DATOS

Recopilación de los datos

En relación con la analítica de los datos se escoge el método Knowledge Discovery in Databases (KDD), el cual describe las fases a seguir para un proceso de minería de datos. Este proceso es complejo e implica no solo la obtención de modelos o patrones, que es el objetivo central de la minería de datos, sino también su evaluación e interpretación (Hernández Orallo et al., 2005). Los sistemas incluyen la selección, limpieza, transformación y proyección de los datos. Además, permiten analizar los datos para identificar patrones y

modelos pertinentes, así como evaluar e interpretar estos patrones para convertirlos en conocimiento útil (Ramos-Serrano, 2021).

En la fase inicial, se procede a la selección de los datos abiertos que han sido reportados en el SIVIGILA) por parte de la alcaldía de Bucaramanga, durante el período comprendido entre los años 2016 y 2023. Este conjunto de datos incluye registros detallados de los casos, abordando variables socioeconómicas, psicológicas y demográficas relevantes; datos que son consumidos a través de la interfaz de programación de aplicación (API) que permite una comunicación directa con el sistema generador. A continuación, en la tabla 1 se presenta una visión general del conjunto de datos.

TABLA 1.
Entidad suministradora de los datos abiertos

Entidad	Enlace	Variable	Formato	Columnas	Filas	Datos
Secretaría de Salud y Medio Ambiente, Alcaldía Bucaramanga	https://www.datos.gov.co/Salud-y-Protecci-n-Social/86-Intencion-suicida-sivigila-enero-2016-a-abril-2/78wv-f29q/about_data https://www.bucaramanga.gov.co/centro_analitica/seriesDatos/index/64	Intento Suicida	CSV/API	51	4175	202758

Fuente: elaboración propia.

Nota: características del Data Set

Preprocesamiento de datos

El preprocesamiento de datos es un paso crucial que abarca técnicas como la limpieza de datos mediante el lenguaje de programación Python desarrollado en el entorno Google Colaboratory. Esta práctica se centra en la eliminación de datos ruidosos y la corrección de posibles inconsistencias en los conjuntos de datos recolectados (Rodríguez González & Ugalde Saborio, 2021). Además de la limpieza, el preprocesamiento implica una serie de pasos adicionales como la integración de datos de múltiples fuentes, la selección de características relevantes y la transformación de datos para su mejor comprensión y análisis (Ramos-Serrano, 2021). Igualmente, implica acciones como estandarizar formatos, codificar caracteres y verificar la integridad para que no haya valores corruptos, entre otros aspectos (Ríos Insua & Gómez-Ullate Oteiza, 2019).

El objetivo principal del preprocesamiento de datos en minería de datos es prepararlos de manera adecuada para su posterior análisis y modelado (Jones, 2019). Esto incluye asegurarse de que los datos estén completos, correctos y coherentes, lo que contribuye a la exactitud, integridad y confiabilidad de los resultados obtenidos durante las etapas posteriores del proyecto (Arriaza Balmón, 2006).

Minería de datos

Al analizar los datos de los casos de intentos suicidas reportados en el Sistema Nacional de Vigilancia en Salud Pública por la alcaldía de Bucaramanga durante los años 2016 y 2023, se identificó un total de 4175 casos en la población.

Se destaca que la mayoría de los casos se registraron en el año 2022, representando el 22,6 % del total, seguido por el año 2021 con un 17,5 %. Además, el año 2019 presentó un 15,6 % de los casos. Es importante señalar que los datos del año 2023 no están actualizados. En relación con la etapa del desarrollo, se observa la siguiente distribución: el 40,5 % de los casos ocurrieron durante la juventud, el 28,1 % durante la adultez, el 26,4 % durante la adolescencia, el 3,3 % en personas mayores, el 1,4 % en la infancia y el 0,05 % en la primera infancia. En la figura 1 se presenta esta información de manera detallada:

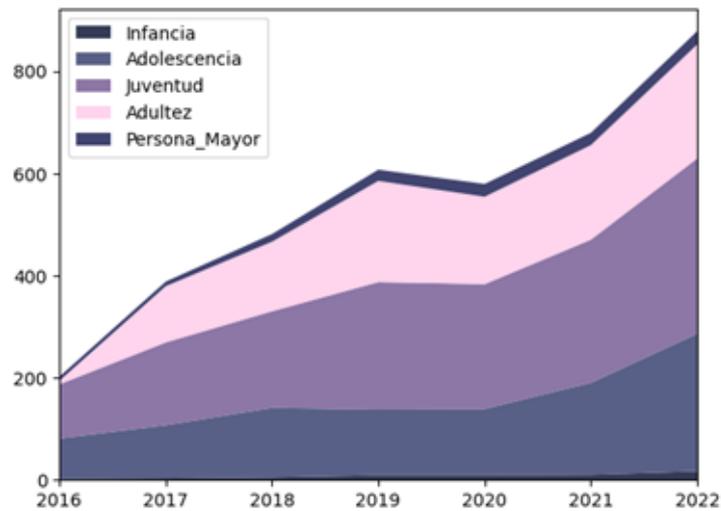


FIGURA 1
Etapas del desarrollo según los casos.

Fuente: elaboración propia, desarrollada en Python y exportada de COLAB

En cuanto al sexo, se observa que los intentos suicidas afectan con mayor frecuencia a las mujeres, representando el 60,5 % de los casos, en comparación con los hombres, que constituyen el 39,5 %. En relación con el régimen al que pertenecen las personas afectadas, se identifica que el 58 % está afiliado al régimen contributivo, el 32,6 % al subsidiado, mientras que un 9,4 % muestra otra forma de afiliación.

Al examinar la zona o área donde se presentaron los casos, se evidencia que el 95 % ocurrió en la cabecera municipal, el 4 % en un centro poblado y el 0,9 % en áreas rurales dispersas. En cuanto a los barrios y municipios fronterizos con mayor incidencia, se destaca que Floridablanca presenta el 14,2 % de los casos, seguido por Girón con el 6,9 %, Campo Hermoso con el 3,7 %, Piedecuesta con el 3 %, y Provenza con el 2,4 %. En la figura 2 se marcan los puntos de ubicación de los sitios con mayor número de casos.

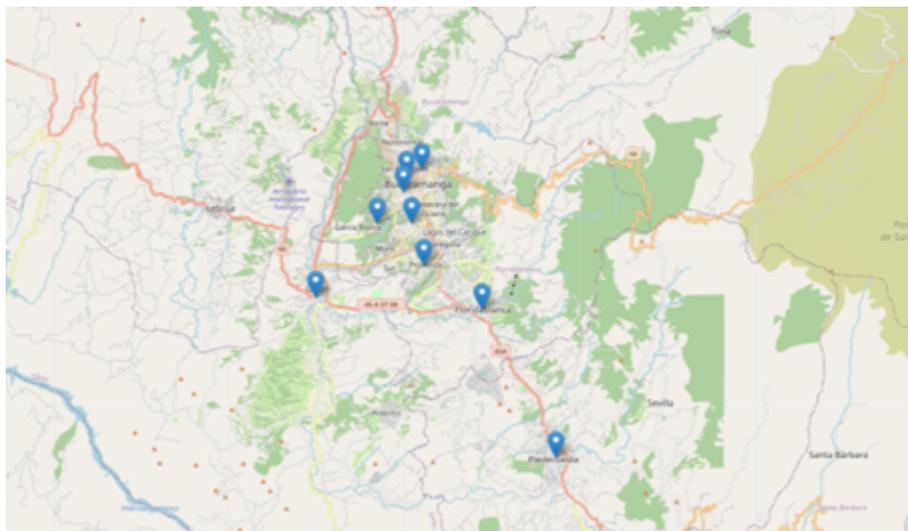


FIGURA 2
Imagen de la ciudad de Bucaramanga-Colombia.

Fuente: elaboración propia, desarrollada en Python y exportada de COLAB.

Se observa que en el 42,1 % de los casos hubo intentos previos, mientras que un 57,8 % no los reporta. Respecto a las ideas suicidas recurrentes, se registra que el 29,1 % las ha experimentado. Asimismo, al indagar sobre la presencia de un plan organizado, el 10,1 % expone haberlo elaborado.

En cuanto a las características socioeconómicas, como se ilustra en la figura 3, un 13,6 % reporta tener problemas económicos y un 5,1 % enfrenta problemas laborales. Igualmente, un 34,9 % expone diferencias con su pareja o expareja y un 3,8 % manifiesta estar involucrado en procesos legales. Además, un 8,1 % presenta problemas escolares o académicos. Por otra parte, se observa que el 15 % de los casos manifestaron problemas familiares, mientras que un 5,7 % tiene antecedentes familiares de conducta suicida. Igualmente, un 3,6 % expone haber experimentado la muerte de un familiar y un 0,9 % indica que un familiar o amigo ha fallecido por suicidio.

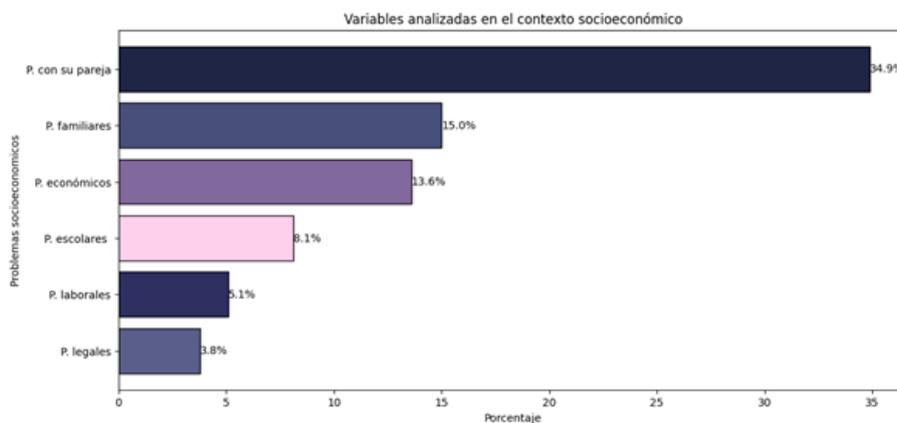


FIGURA 3
VARIABLES ANALIZADAS EN EL CONTEXTO SOCIOECONÓMICO.
 Fuente: elaboración propia, desarrollada en Python y exportada de COLAB.

En cuanto a la salud, como se evidencia en la figura 4, el 7 % indica padecer una enfermedad crónica y un 7 % ha sido víctima de maltrato físico, psicológico o sexual. Sobre el consumo de sustancias psicoactivas, el 20,9 % manifiesta ser consumidor, mientras que el 6,4 % admite abusar del alcohol. En relación con la salud mental, se identifica que el 43,6 % tiene antecedentes psiquiátricos, el 38,3 % manifiesta trastorno depresivo, el 2,8 % padece trastorno de personalidad, el 5,5 % tiene trastorno bipolar y el 2,6 % padece esquizofrenia.

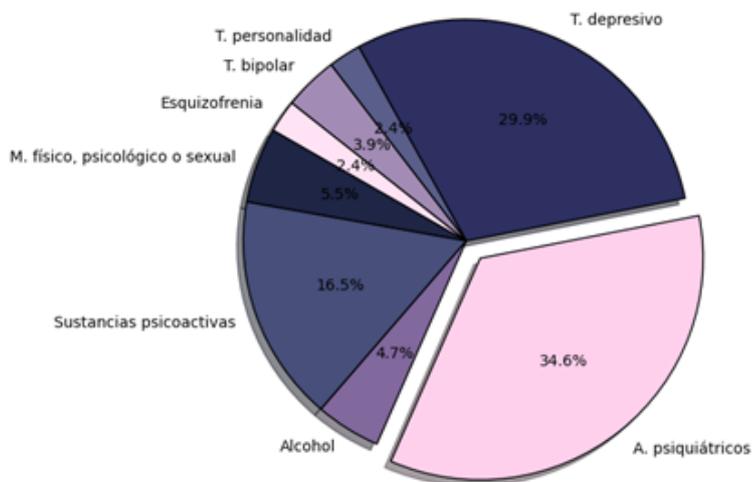


FIGURA 4
 Variables analizadas en relación con la salud mental.
 Fuente: elaboración propia, desarrollada en Python y exportada de COLAB

En cuanto a la forma de los intentos suicidas, el 5,3 % optó por el ahorcamiento, el 24,3 % utilizó armas cortopunzantes, el 0,4 % intentó con arma de fuego, el 0,1 % con la inmolación, el 6,7 % se lanzó al vacío, el 1,3 % optó por lanzarse a un vehículo y un 0,1 % se lanzó al agua. El 36,1 % utilizó la intoxicación como método, siendo los medicamentos, plaguicidas y las sustancias psicoactivas los más comunes. Finalmente, estas conductas suicidas desencadenaron en un 17,1 % de los casos la hospitalización.

Como se muestra en la figura 5, en cuanto a la relación entre la variable sexo y otras variables significativas, se observa que los hombres tienen un 43,4 % más de intentos previos que las mujeres, quienes presentan un 41,3 %. Asimismo, el grupo de hombres exhibe un 30,3 % de ideas suicidas persistentes, en comparación con el 28,3 % registrado en las mujeres. Además, los hombres muestran un 33,7 % de consumo de sustancias psicoactivas.

Por otro lado, las mujeres presentan un mayor número de casos de antecedentes psiquiátricos, con un 44,8 %, mientras que los hombres muestran un 41,9 %. Igualmente, las mujeres muestran una mayor incidencia de trastornos depresivos, con un 40,9 %, en comparación con el 34,3 % de los hombres. Además, las mujeres tienen un 36,6 % más casos de problemas de pareja o expareja que los hombres, quienes registran un 32,3 %. Es importante destacar que el grupo de mujeres presenta un 70,8 % de intoxicaciones en comparación con el 53 % de los hombres, lo que sugiere una diferencia significativa en este aspecto entre los dos grupos de género.

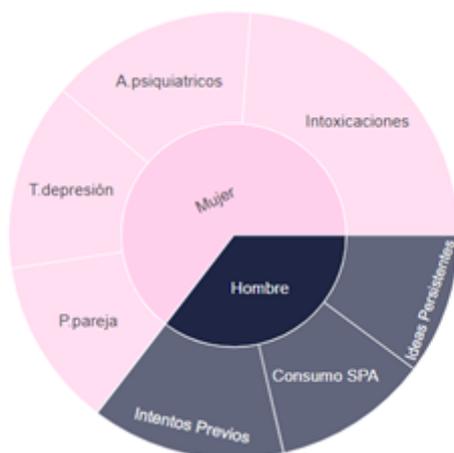


FIGURA 5
 Variables analizadas en relación con el sexo.
 Fuente: elaboración propia, desarrollada en Python y exportada de COLAB.

Análisis de los datos de los intentos suicidas mediante el modelo de Random Forest

En cuanto al análisis de los datos, se emplea el algoritmo de Random Forest como herramienta altamente versátil en el campo del aprendizaje automático. Este método no solo realiza técnicas de reducción dimensional, aborda valores faltantes y atípicos, y otros procesos esenciales de exploración de datos, sino que también sobresale en su desempeño. Este enfoque se basa en el aprendizaje por conjuntos, donde se combinan varios modelos más simples para formar uno más robusto (González, 2019).

La práctica común consiste en generar múltiples versiones del conjunto de entrenamiento mediante muestreo con reemplazo (*bagging*). Durante la construcción de cada árbol de decisión, se elige al azar un subconjunto de variables del conjunto de datos. Esto permite evaluar la importancia relativa de cada variable al estimar el error del clasificador combinado cuando se altera una variable y se permutan sus valores en el conjunto de pruebas. Comparar este error con el error sin permutación ayuda a determinar la relevancia de cada variable (Gironés Roig et al., 2017).

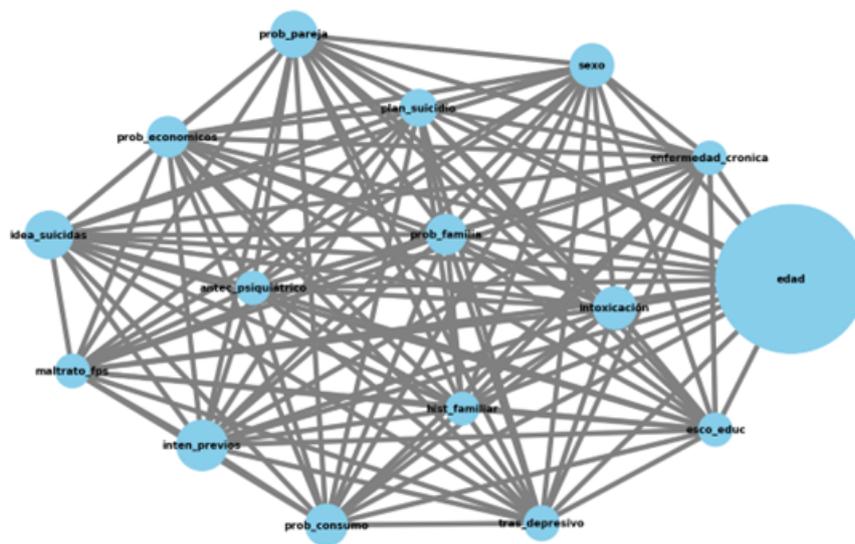


FIGURA 6
Relevancia de las variables.

Fuente: elaboración propia, desarrollada en Python y exportada de COLAB.

Nota: la estructura de grafos muestra la relevancia de cada variable al aplicar el modelo a los datos.

La figura 6 expone una red de nodos conectados entre sí, con palabras dentro de cada círculo que representan la variable. Las palabras muestran diferentes factores que pueden estar asociados con el intento de suicidio. Además, los círculos están conectados por aristas, lo que indica una relación entre los diferentes factores. Igualmente, el modelo de Random Forest posee la habilidad de identificar las variables que resultan relevantes o que contribuyen de manera significativa a los modelos de clasificación o predicción. En consecuencia, se evidencia que las variables más representativas son la edad, el género, los intentos previos, la persistencia de la ideación suicida, los problemas de pareja, el consumo de sustancias, así como los problemas económicos.

Análisis de clasificación binaria de los casos de intentos suicidas mediante modelos de machine learning

Se realiza un análisis de clasificación de casos que requirieron hospitalización, mediante el empleo de algoritmos de *machine learning*, con el objetivo de ofrecer una contribución a la mejora de la gestión de la salud pública. La ampliación de este estudio a través de la inclusión de una variedad de algoritmos y técnicas de preprocesamiento de datos podría conducir a una mejora sustancial en la precisión y utilidad de los resultados obtenidos. Este enfoque permitiría obtener una comprensión profunda de los factores determinantes que inciden en la hospitalización, lo que a su vez facilitaría la implementación de intervenciones efectivas.

EVALUACIÓN DE LOS MODELOS

El análisis se enfoca en evaluar métricas clave como la exactitud, el *recall* y el F1 Score en los algoritmos de *machine learning* para clasificar los casos de personas que necesitaron hospitalización. La exactitud refleja la proporción de predicciones correctas, pero el *recall* y el F1 Score son vitales para detectar correctamente los casos que requieren hospitalización, especialmente en situaciones de desequilibrio de clases. El *recall* mide la capacidad del modelo para identificar estos casos, mientras que el F1 Score combina precisión y *recall* para evaluar la capacidad general del modelo. Estas métricas permiten una evaluación exhaustiva del rendimiento

del algoritmo en la predicción de casos de personas hospitalizadas, siendo esenciales para su aplicación en el ámbito de la salud. En la tabla 2 se muestran los resultados de esta evaluación.

TABLA 2.
Evaluación de los algoritmos de machine learning aplicados a los datos

Algoritmo	Accuracy	F1 Score	Recall
XGBoost	0,8298969072164949	0,7536866560185859	0,8298969072164949
Redes Neuronales	0,830927848815918	0,9076576576576576	1,0
LTSM			
Regresión Logística	0,8309278350515464	0,9076576576576576	1,0
Support Vector Machine (SVM)	0,8350515463917526	0,9101123595505618	1,0
K-Nearest Neighbors (KNN)	0,8092783505154639	0,8938307030129125	0,978021978021978

Fuente: elaboración propia.

Nota: algoritmos desarrollados en lenguaje Python y aplicados a los datos al clasificar casos de hospitalización.

De forma general, el modelo Support Vector Machine (SVM) es identificado como el mejor en este caso, lo que lo convierte en una buena opción si la precisión es la principal prioridad, lo que indica su capacidad para identificar verdaderos positivos y minimizar los falsos positivos, aspecto crítico en muchas aplicaciones. Además, los SVM son robustos y pueden manejar conjuntos de datos complejos, incluso cuando las características son muchas en comparación con las muestras. Su flexibilidad, gracias al uso de diferentes kernels, permite capturar relaciones no lineales entre las características.

DISCUSIÓN

Este estudio ofrece una contribución significativa al objetivo establecido al abordar el intento suicida como una fase específica o una manifestación dentro del comportamiento suicida, el cual se encuentra intrínsecamente ligado a otras variables fundamentales. Tras el análisis de los resultados de este estudio, se revela que la juventud emerge como el grupo demográfico con la mayor prevalencia de intentos de suicidio, representando un 40,5 % del total de casos. Este hallazgo corrobora las observaciones realizadas por Restrepo Tobón et al. (2023), cuyo estudio concluye que los problemas psicosociales afectan de manera más significativa a los individuos más jóvenes. Por lo tanto, se evidencia que los jóvenes son más vulnerables al comportamiento suicida.

Los datos indican que la mayoría de los incidentes ocurrieron en el año 2022, representando el 22,6 % del total, seguido por el año 2021 con un 17,5 %. Asimismo, el año 2019 registró un 15,6 % de los casos. Según los datos del Ministerio de Salud y Protección Social de Colombia (2021), en el período comprendido entre 2017 y 2020, el año con el mayor número de casos de muertes e intentos de suicidio en jóvenes de 15 a 19 años fue 2020. Este aumento, según la entidad, se atribuye a la persistencia de la pandemia, la cual incremento el riesgo de suicidio, especialmente entre aquellos que sufren de trastornos mentales. Esto se debe al temor, la soledad, el distanciamiento físico y otros factores estresantes, como el estigma y las dificultades financieras.

Al analizar la variable sexo, se destaca que las mujeres representan el 60,5 % de los casos afectados, en comparación con los hombres. Asimismo, se observa que el 95 % de los casos ocurren en áreas urbanas, y la mayoría de las personas afectadas pertenecen al régimen contributivo de salud. Estas tendencias son

corroboradas por el estudio de Vigoya-Rubiano et al. (2019). Sin embargo, difiere en cuanto al régimen de salud predominante en su estudio, el cual corresponde al régimen subsidiado, debido a las características territoriales y cantidad de casos en estudio.

Los hallazgos también enfatizan en la relevancia de considerar los antecedentes, los intentos previos de suicidio, que se registran en un 42,1 % de los casos, así como las tensiones con la pareja o expareja, que constituyen el 34,9 %, y el uso de sustancias psicoactivas, con un 20,9 %. Estas situaciones han sido consistentemente identificadas en diversas investigaciones, resaltando la complejidad de los factores que contribuyen al riesgo de suicidio (Campillo Serrano & Fajardo Dolci, 2021; Cañón Buitrago & Carmona Parra, 2018; González Sancho & Picado Cortés, 2020; Santillán Pizarro & Pereyra, 2021; Soto Salcedo et al., 2020). Se subraya que por cada suicidio consumado, se han documentado varios intentos de suicidio (Del Cid Rodríguez, 2020). Además, en lo individual, se señalan como factores que aumentan este riesgo la desesperanza, el abuso de sustancias, la presencia de ciertas enfermedades, el historial de intentos de suicidio previos y las experiencias de victimización o perpetración de violencia (Soto Salcedo et al., 2020).

Por otra parte, un punto destacado en relación con la salud mental se centra en que el 43,6 % de los sujetos presenta antecedentes psiquiátricos, mientras que el 38,3 % manifiesta trastorno depresivo. Esta evidencia subraya la estrecha relación entre los trastornos mentales y los intentos suicidas (Amaral et al., 2020; Pabon, 2021; Sindeev et al., 2019). En cuanto a la modalidad de los intentos suicidas, se observa que el 24,3 % de los casos implicaron el uso de armas cortopunzantes, mientras que el 36,1 % empleó la intoxicación como método. Estas conductas también han sido documentadas en investigaciones previas, que han señalado la intoxicación como el método más común para llevar a cabo intentos de suicidio (Beitia-Cardona et al., 2019; Calderón-Ramírez et al., 2017; Cortés Alfaro et al., 2019; López-Steinmetz, 2019). Además, factores como el acoso y la vulnerabilidad a diversas formas de violencia, junto con la exposición a condiciones socioeconómicas precarias, indican un mayor riesgo de intentos de suicidio (Santillán Pizarro & Pereyra, 2021).

Al abordar cualquier análisis o aplicación de modelos de *machine learning*, es crucial destacar las limitaciones inherentes a la calidad y cantidad de los datos disponibles. Estas limitaciones pueden ejercer una influencia significativa en los resultados obtenidos. Hay que considerar la posibilidad de sesgos en los datos, la falta de representatividad de la población objetivo, errores de medición y la falta de datos relevantes para ciertos aspectos del problema en estudio. Por otra parte, la selección inadecuada de características o una alta dimensionalidad de los datos pueden conducir a un sobreajuste o subajuste del modelo, lo que afectaría la generalización de los resultados obtenidos.

CONCLUSIONES

Los resultados muestran una concentración notable de casos en 2022, seguido de 2021 y 2019, indicando posibles cambios en la detección o factores desencadenantes. En cuanto a las etapas del desarrollo, se observa un predominio de casos durante la juventud, seguida por la adolescencia. La falta de datos actualizados para 2023 limita las comparaciones. Estos hallazgos resaltan la importancia de una vigilancia continua y de estrategias específicas para abordar los diferentes períodos temporales y grupos de edad afectados.

Por otro lado, se revela una prevalencia significativa de intentos suicidas entre las mujeres, con una mayor afiliación al régimen contributivo y una concentración en áreas urbanas. Los casos previos de intentos suicidas son comunes, al igual que las ideas suicidas recurrentes. Los factores socioeconómicos subyacentes incluyen problemas familiares y conflictos con la pareja o expareja. El consumo de sustancias psicoactivas se destaca como un factor relevante. Además, los antecedentes psiquiátricos y los trastornos depresivos son comunes entre los afectados. Los métodos utilizados en los intentos incluyen el uso de armas cortopunzantes y la intoxicación, particularmente con medicamentos y plaguicidas.

En cuanto a las diferencias de género, se observa una mayor proporción de intentos previos, ideas suicidas persistentes y consumo de sustancias psicoactivas entre los hombres. Por otro lado, las mujeres muestran una mayor prevalencia de antecedentes psiquiátricos, trastornos depresivos, problemas de pareja o expareja, y casos de intoxicación.

El método Random Forest demostró su utilidad al identificar factores relevantes asociados con intentos suicidas. Su eficacia se resalta por la capacidad para explorar datos complejos y precisamente identificar variables significativas como edad, género, intentos previos, ideación suicida persistente, problemas de pareja, consumo de sustancias y problemas económicos. Este enfoque ofrece una comprensión profunda de las interacciones entre estos factores, proporcionando una base sólida para estrategias de prevención e intervención efectivas.

El uso de algoritmos de *machine learning* para el análisis y clasificación de casos de hospitalización con el propósito de identificar sus características representa una estrategia innovadora. La evaluación se enfoca en métricas fundamentales como la exactitud, el *recall* y el F1 Score. Entre los modelos evaluados, el Support Vector Machine (SVM) ha demostrado ser el más eficaz en este contexto, gracias a su capacidad para minimizar los falsos positivos y gestionar conjuntos de datos complejos. Estos resultados sugieren que el algoritmo es adecuado para aplicaciones en el ámbito de la salud.

Para abordar el problema de los intentos suicidas, es fundamental implementar programas preventivos dirigidos a grupos de alto riesgo y comunidades vulnerables, así como desarrollar intervenciones específicas centradas en factores socioeconómicos y de salud mental. Además, la aplicación continua de técnicas de *machine learning* en la clasificación de casos de hospitalización puede mejorar la detección temprana y la asignación de recursos para una atención más efectiva, mientras que la integración de tecnología para el seguimiento y apoyo a largo plazo puede proporcionar intervenciones personalizadas y oportunas fuera del entorno hospitalario.

REFERENCIAS

- Al-Halabí, S., & Fonseca-Pedrero, E. (coords.) (2023). *Manual de psicología de la conducta suicida*. Piramide.
- Amaral, A. P., Uchoa Sampaio, J., Ney Matos, F. R., Santos Pocinho, M. T., Fernandes de Mesquita, R., & Milanês Sousa, L. R. (2020). Depresión e ideación suicida en la adolescencia: Implementación y evaluación de un programa de intervención. *Enfermería Global*, 19(59), 1-35. <https://doi.org/10.6018/eglobal.402951>
- Arriaza Balmón, M. (2006). *Guía práctica de análisis de datos*. Instituto de Investigación y Formación Agraria y Pesquera.
- Avila Pérez, M. L. (2021). *Modelo de Predicción de Deserción Estudiantil, Apoyado en Tecnologías de Data Mining, en un Curso de Primera Matrícula de la Escuela ECBTI de la UNAD* [tesis de maestría, Universidad Nacional Abierta y a Distancia]. Repositorio UNAD. <https://repository.unad.edu.co/bitstream/handle/10596/42544/Mavila.pdf?sequence=1&isAllowed=y>
- Baena Paz, G. (2014). *Metodología de la investigación*. Grupo Editorial Patria.
- 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>
- Biscaia Fernández, J. M., Mohedano del Pozo, R. B., & Biscaia Fernández, C. J. (2023). La inteligencia artificial en la prevención de conductas suicidas: Aspectos técnicos y consideraciones ético-legales. *Revista de Bioética y Derecho*, (59), 181-203. <https://doi.org/10.1344/rbd2023.59.42759>
- 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), 149-159. <https://doi.org/10.21676/2389783X.1963>
- Caminotti, M., & Toppi, H. P. (2020). *Metodología de la investigación social: Caja de herramientas*. Eudeba.

- 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>
- Cañón Buitrago, S. C., & Carmona Parra, J. A. (2018). Ideación y conductas suicidas en adolescentes y jóvenes. *Pediatría Atención Primaria*, 20(80), 387-397. http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1139-76322018000400014&lng=es.
- Castillo-Zúñiga, I., Lunas-Rojas, F. J., & López-Veyna, J. I. (2022). Detección de rasgos en estudiantes con tendencia suicida en Internet aplicando Minería Web. *Comunicar*, 30(71), 105-117. <https://doi.org/10.3916/C71-2022-08>
- Castro-Osorio, R., Maldonado-Avendaño, N., & Cardona-Gómez, P. (2022). Propuesta de un modelo de la ideación suicida en estudiantes de Medicina en Colombia: Un estudio de simulación. *Revista Colombiana de Psiquiatría*, 51(1), 17-24. <https://doi.org/10.1016/j.rcp.2020.09.001>
- Cortés Alfaro, A., Suárez Medina, R., & Serra Larín, S. (2019). Métodos y sustancias empleados en la conducta suicida en adolescentes. *Revista Cubana de Medicina General Integral*, 35(4), 1-15. <https://revmgi.sld.cu/index.php/mgi/article/view/1105>
- Del Cid Rodríguez, A. (2020). Intento suicida en adolescentes: un abordaje desde una perspectiva cultural en las familias. *Enfoque*, 27(23), 25-45. <https://revistas.up.ac.pa/index.php/enfoque/article/view/2234>
- Domínguez Granda, J. B. (2015). *Manual de metodología de la investigación científica*. Universidad Católica Los Angeles de Chimbóte.
- Gironés Roig, J., Casas Roma, J., Minguillón Alfonso, J., & Caihuélas Quiles, R. (2017). *Minería de datos Modelos y algoritmos*. Editorial UOC.
- González, L. (2019). *Machine Learning con Python Aprendizaje Supervisado*. www.detodopython.com.
- González Sancho, R. D., & Picado Cortés, M. (2020). Revisión sistemática de literatura sobre suicidio: Factores de riesgo y protectores en jóvenes latinoamericanos 1995-2017. *Actualidades en Psicología*, 34(129), 47-69. <https://doi.org/10.15517/ap.v34i129.34298>
- Hernández Orallo, J., Ramírez Quintana, M. J., & Ferri Ramírez, C. (2005). *Introducción a la minería de datos*. Pearson Educación.
- Hernández Sampieri, R., Mendoza Torres, P., Méndez Valencia, S., & Cuevas Romo, A. (2019). *Metodología de la investigación para bachillerato*. Mc Graw Hill.
- Hinojosa-Vega, R., Rojas-Carrasco, K. E., González-Muñiz, S., Jaramillo-Villanueva, L., & Jiménez-Márquez, A. (2023). Intento de suicidio en paciente escolar y factores relacionados: Reporte de caso. *Revista Médica del Instituto Mexicano del Seguro Social*, 61(6), 875-881. <https://doi.org/10.5281/zenodo.10064736>
- Jones, H. (2019). *Minería de Datos Guía de Minería de Datos para Principiantes, que Incluye Aplicaciones para Negocios, Técnicas de Minería de Datos, Conceptos y Más*. Ch Publications.
- Joyanes Aguilar, L. (2019). *Inteligencia de negocios y analítica de datos. Una visión global de Business Intelligence & Analytics*. Alfaomega.
- López Steinmetz, L. C. (2019). Perfil de métodos en intentos de suicidio: Tendencias e implicancias para la prevención. Jujuy, noroeste de Argentina. *Ciencias Psicológicas*, 13(2), 197-208. <https://doi.org/10.22235/cp.v13i2.1872>
- Ministerio de Salud y Protección Social de Colombia. (2018). *Boletín de salud mental Conducta suicida Subdirección de Enfermedades No Transmisibles*. Minsalud. <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/ENT/boletin-conducta-suicida.pdf>
- Ministerio de Salud y Protección Social de Colombia. (2021). *Estrategia Nacional para la Prevención de la Conducta Suicida en Colombia*. Minsalud. <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/ENT/estrategia-nacional-conducta-suicida-2021.pdf>
- Mojica Walteros, C. M., Hoyos Gómez, L. K., Vanegas Gama, H. S., Muñoz Torres, L. D., & Fernández-Ávila, D. G. (2023). Intento de suicidio pediátrico e ingreso a Unidad de Cuidado Intensivos, antes y después de la pandemia, en un hospital universitario en Boyacá, Colombia. *Pediatría*, 56(2), 120-135. <https://doi.org/10.14295/rp.v56i2.388>

- Morfin López, T., & Ibarra López, A. M. (2018). *Fenómeno suicida: Un acercamiento transdisciplinar*. Manual Moderno.
- Mosquera Bahamón, Y. (2020). *Intento de suicidio Colombia 2020*. Instituto Nacional de Salud. https://www.ins.gov.co/buscador-eventos/Informesdeevento/INTENTO%20DE%20SUCIDIO_2020.pdf
- Myers, D. G., & Twenge, J. M. (2019). *Psicología Social*. Mc Graw Hill.
- Navarro Acosta, J. A., Soto Mendoza, V., Saucedo Zendejo, F. R., Guajardo Espinoza, J. M., & Rivera Morales, M. T. (2021). Evaluación psicológica de profesores y alumnos mexicanos durante la pandemia de COVID-19 mediante técnicas de Machine learning. *Ingeniería, investigación y tecnología*, 22(4), 1-20. <https://doi.org/10.22201/fi.25940732e.2021.22.4.026>
- Organización Mundial de la Salud. (2021, 17 de junio). *Suicidio*. Centro de prensa OMS. <https://www.who.int/es/news-room/fact-sheets/detail/suicide>
- Pabon, A. Y. (2021). Intentos de suicidio y Trastornos mentales. *Revista Habanera de Ciencias Médicas*, 20(4), 20-40. <https://revhabanera.sld.cu/index.php/rhab/article/view/3967>
- Papalia, D. E., & Martonell, G. (2017). *Desarrollo humano*. Mc Graw Hill.
- Paredes Lascano, P. L., Bravo Paredes, L. A., Posso Lopez, G. P., & Quelal Zurita, A. D. (2022). Factores asociados al intento de suicidio en adolescentes. *Revista Médica-Científica CAMBIOS HECAM*, 21(1), 115-135. <https://doi.org/10.36015/cambios.v21.n1.2022.848>
- Ramos-Serrano, C. (2021). *Aprende minería de datos con modelos y algoritmos: Aprende de minería de datos aplicado al análisis predictivo*. Carmelo Ramos.
- Restrepo Tobón, M. V., Mejía Zapata, S. I., & García-Peña, J. J. (2023). El intento suicida en relación con la edad y otras variables psicosociales. Un análisis de contexto. *Revista Virtual Universidad Católica del Norte*, (69), 300-315. <https://doi.org/10.35575/rvucn.n69a9>
- Ríos Insua, D., & Gómez-Ullate Oteiza, D. (2019). *Big data Conceptos, tecnologías y aplicaciones*. Editorial CSIC.
- Robles Sánchez, J. I., Acinas Acinas, M. P., & Pérez-Íñigo Gancedo, J. L. (2020). *Estrategias de intervención psicológica en la conducta suicida*. Editorial Síntesis.
- Rodríguez González, J., & Ugalde Saborio, E. (2021). Impacto de la estandarización y escalado: Factor para predicción de costos en proyectos a través de una red neuronal artificial. *Ingeniare. Revista chilena de ingeniería*, 29(2), 265-275. <https://doi.org/10.4067/S0718-33052021000200265>
- Santillán Pizarro, M. M., & Pereyra, E. J. (2021). Factores asociados a la ideación del suicidio en adolescentes. Algunas evidencias en Argentina. *Astrolabio*, 34(26), 1-25. <https://doi.org/10.55441/1668.7515.n26.27659>
- Sindeev, A., Arispe Alburqueque, C. M., & Villegas Escarate, J. N. (2019). Factores asociados al riesgo e intento suicida en estudiantes de medicina de una universidad privada de Lima. *Revista Medica Herediana*, 30(4), 232-241. <https://doi.org/10.20453/rmh.v30i4.3658>
- Soto Salcedo, A., Villaroel Grüner, P., & Véliz Burgos, A. (2020). Factores que intervienen en riesgo suicida y parasuicida en jóvenes Chilenos. *Propósitos y Representaciones*, 8(3), 1-20. <https://doi.org/10.20511/pyr2020.v8n3.672>
- Vigoya-Rubiano, A. R., Osorio-Barajas, Y. S., & Salamanca-Camargo, Y. (2019). Caracterización del intento de suicidio en una ciudad colombiana (2012–2017). *Revista Duazary*, 16(1), 53-66.