



Revista Científica General José María Córdova

ISSN: 1900-6586

ISSN: 2500-7645

Escuela Militar de Cadetes "General José María Córdova"

Fontalvo Herrera, Tomás José; Vega Hernández, María Alejandra; Mejía Zambrano, Fabio  
Método de *clustering* e inteligencia artificial para clasificar y proyectar delitos violentos en Colombia  
Revista Científica General José María Córdova, vol. 21, núm. 42, 2023, Abril-Junio, pp. 550-572  
Escuela Militar de Cadetes "General José María Córdova"

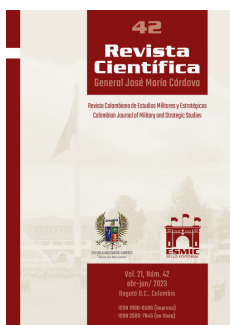
DOI: <https://doi.org/10.21830/19006586.1117>

Disponible en: <https://www.redalyc.org/articulo.oa?id=476277508011>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

UAEH  
redalyc.org

Sistema de Información Científica Redalyc  
Red de Revistas Científicas de América Latina y el Caribe, España y Portugal  
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso  
abierto



## Revista Científica General José María Córdova

(Revista Colombiana de Estudios Militares y Estratégicos)

Bogotá D.C., Colombia

ISSN 1900-6586 (impreso), 2500-7645 (en línea)

Web oficial: <https://www.revistacientificaesmic.com>

## Método de *clustering* e inteligencia artificial para clasificar y proyectar delitos violentos en Colombia

**Tomás José Fontalvo Herrera**

<https://orcid.org/0000-0003-4642-9251>

[tfontalvoh@unicartagena.edu.co](mailto:tfontalvoh@unicartagena.edu.co)

Universidad de Cartagena, Colombia

**María Alejandra Vega Hernández**

<https://orcid.org/0000-0001-5898-7440>

[mvegahl@unicartagena.edu.co](mailto:mvegahl@unicartagena.edu.co)

Universidad de Cartagena, Colombia

**Fabio Mejía Zambrano**

<https://orcid.org/0000-0002-6689-3345>

[fmejiaz@ut.edu.co](mailto:fmejiaz@ut.edu.co)

Universidad del Tolima, Ibagué, Colombia

**Citación APA:** Fontalvo Herrera, T. J., Vega Hernández, M. A., & Mejía Zambrano, F. (2023). Método de *clustering* e inteligencia artificial para clasificar y proyectar delitos violentos en Colombia.. *Revista Científica General José María Córdova*, 21(42), 551-572. <https://dx.doi.org/10.21830/19006586.1117>

**Publicado en línea:** 1.º de abril de 2023

Los artículos publicados por la *Revista Científica General José María Córdova* son de acceso abierto bajo una licencia Creative Commons: Atribución - No Comercial - Sin Derivados.



**Para enviar un artículo:**

<https://www.revistacientificaesmic.com/index.php/esmic/about/submissions>



Miles Doctus

## Método de *clustering* e inteligencia artificial para clasificar y proyectar delitos violentos en Colombia

### Clustering method and artificial intelligence to classify and project violent crime in Colombia

**Tomás José Fontalvo Herrera y María Alejandra Vega Hernández**

Universidad de Cartagena, Colombia

**Fabio Mejía Zambrano**

Universidad del Tolima, Ibagué, Colombia

**RESUMEN.** El objetivo de esta investigación es proponer clústeres de delitos violentos en Colombia por departamentos junto con una estructura de redes neuronales para su clasificación y pronóstico. Para ello, se parte del análisis del método de *clustering*, la inteligencia artificial y la definición de delitos violentos. Con un método explicativo cuantitativo, se partió de los datos generados por la Policía Nacional sobre delitos entre 2018 y 2022. Como resultado, se establecieron cuatro clústeres de delitos y factores de violencia que caracterizan grupos de departamentos, lo que permitió identificar regiones con mayor y menor impacto de actos delictivos. Luego se planteó una red neuronal de doble capa que alcanzó una capacidad de clasificación y predicción de 97,7 % para los delitos según su tipo e impacto, lo cual espera servir como herramienta para la prevención de delitos.

**PALABRAS CLAVE:** análisis de clúster; delitos violentos; inteligencia artificial; modelo de simulación; redes neuronales; violencia

**ABSTRACT.** The objective of this research is to propose clusters of violent crimes in Colombia by departments together with a neural network structure for their classification and prognosis. For this purpose, it is based on the analysis of the clustering method, artificial intelligence and the definition of violent crimes. Using a quantitative explanatory method, the data generated by the National Police on crimes between 2018 and 2022 was used. As a result, four clusters of crimes and violence factors that characterize groups of departments were established, which allowed identifying regions with higher and lower impact of criminal acts. A double-layer neural network was then proposed that achieved a classification and prediction capacity of 97.7% for crimes according to their type and impact, which is expected to serve as a tool for crime prevention.

**KEYWORDS:** artificial intelligence; cluster analysis; neural networks; simulation model; violence; violent crimes

Sección: DOSIER • Artículo de investigación científica y tecnológica

Recibido: 2 de diciembre de 2022 • Aceptado: 7 de marzo de 2023

**CONTACTO:** Fabio Mejía Zambrano ✉ [fmejiaz@ut.edu.co](mailto:fmejiaz@ut.edu.co)

## Introducción

Existen diferentes factores que permiten identificar los países que cuentan con mayores niveles de bienestar social, entendido este como el conjunto de condiciones que se miden a través de índices de calidad de vida: el trabajo, los ingresos, el ambiente, la educación, la salud, entre otros índices, que orientan la construcción de sociedades con los más altos estándares sociales, como lo evidencia la Organización para la Cooperación y el Desarrollo Económicos (OECD, 2021). Entre todos estos factores, sobresale la percepción del estado de tranquilidad, protección social y bienestar de cada país, que se asocia con el indicador de seguridad como causa del nivel de vida (OECD, 2021).

La seguridad ha sido objeto de estudio ante los constantes hechos de conflicto y violencia vividos históricamente. A raíz de esto, se ha establecido una relación entre seguridad y desarrollo (Duffield, 2020). Países de América Latina como Guatemala, Honduras, Nicaragua, Colombia y Venezuela han tenido altos índices de desigualdad, homicidios, pobreza, migración y conflicto, que se han justificado por falta de acciones de los Estados y el consecuente desbalance de los aspectos sociales que construyen el bienestar social. La Unión Europea recoge estos factores de bienestar social como un indicativo de desarrollo y seguridad, de forma que se pueda establecer la disminución de actos violentos y los avances en materia de seguridad (Gratius & Rodríguez, 2021).

El concepto de *seguridad*, según Kahhat (2019), se concibe desde diferentes enfoques dependiendo de sus fines, amenazas, objetos de referencia y medios. En este sentido, cabe comprender la seguridad como la carencia de amenazas; y en relación con los Objetivos de Desarrollo Sostenible, la seguridad implica una lucha contra la desigualdad, los actos delictivos y la violencia, así como el refuerzo de políticas encaminadas a disminuir la inseguridad (Duffield, 2020; Gratius & Rodríguez, 2021).

Los factores de violencia entorpecen el progreso y avance de la seguridad social. Respecto a ellos, se han estudiado y establecido diferentes aspectos (Valencia, 2018; Niño, 2020; Rodríguez et al., 2020). Así, hay factores sociales como el conflicto armado, el desplazamiento que este genera y la sobrepoblación en países desiguales e inequitativos, como puede medirse con el coeficiente de Gini (Datosmacro, 2021), que muestra una alta desigualdad en la mayoría de los países estudiados y un incremento de los índices de inseguridad y delitos cometidos, dada la falta de empleo y la complejidad de conseguir sustento para suplir las necesidades básicas. Asimismo, la escasa educación de calidad, la desigualdad social y la falta de acción del Estado da pie al surgimiento de actos criminales y de violencia (Palacio, 2020).

América Latina es considerada un territorio pacífico, por cuanto no hay conflictos entre naciones, pero paradójicamente lleno de violencia. Un claro ejemplo es Colombia, un país en el que el conflicto armado se ha debilitado, pero ha perdurado hasta hoy en día (Niño, 2020), por lo cual los delitos de violencia tienen una incidencia importante, que difiere en cada región del país. Por eso es importante caracterizar estos delitos para poder tomar acciones en pro de mitigarlos, según el contexto específico en el que ocurran.

Colombia ha vivido un conflicto armado por más de sesenta años cuyas consecuencias y dinámicas de violencia persisten en la actualidad, como se observa en las cifras de la OECD (2021). Así, Colombia obtuvo para 2021 un 43,24 % en porcentaje de seguridad tanto rural como urbana, en contraste con Noruega, el país de mayor índice de bienestar, con un porcentaje de 93,08 %. Esto evidencia una gran diferencia, producto de la inestabilidad en seguridad y bienestar para la población colombiana. Esta problemática hace necesarias las investigaciones sobre las causas, el control y el abordaje de los factores de violencia en el país.

En ese sentido, considerando la percepción de la ciudadanía, la Encuesta de Convivencia y Seguridad Ciudadana (ECSC), aplicada por el Departamento Administrativo Nacional de Estadística (DANE, 2021), muestra que la percepción de inseguridad asciende, luego de una bajada en 2020, de 39 % a 44 %, sobrepasando la tasa de inseguridad de 2019 (39,7 %), mientras que, para el área urbana y rural, la tasa de inseguridad aumentó en un 6,2 % y 0,3 % respectivamente.

Asimismo, el aumento para septiembre de 2022 de algunos delitos violentos, como la extorsión, las amenazas, el secuestro y el terrorismo, en comparación con 2021 (Policía Nacional de Colombia, 2022), reafirman el incremento de la inseguridad percibido por la ciudadanía, y que, en consecuencia, reduce el bienestar social. Esto conlleva nuevos retos para las autoridades encargadas de garantizar la seguridad pública de Colombia —en este caso, la Policía Nacional de Colombia—, que requiere tomar acciones para atender de manera rápida y eficaz las situaciones problemáticas, de modo que disminuya la ocurrencia de delitos en las zonas críticas de Colombia, en aras de la tranquilidad y bienestar de la población.

La Policía Nacional de Colombia ha adelantado operativos para controlar hechos delictivos a través de la instalación de sistemas de vigilancia, establecimiento de cuadrantes, actividades de sensibilización ciudadana y registros de delitos en bases de datos, clasificados por su tipo, para mayor entendimiento y análisis (Ordóñez et al., 2020).

Sin embargo, pese a que estas herramientas reducen los delitos y posibilitan un accionar oportuno, el abordaje pleno y pertinente de esta problemática requiere una mayor aplicabilidad de técnicas inteligentes que converjan en la celeridad y el combate de delitos. Así lo argumenta Tello (2022), al mencionar que dichos mecanismos contribuyen a la evaluación y aplicación de estrategias que inciden en la validez y eficiencia faltante. En consecuencia, esta investigación, basada en inteligencia artificial, junto con el tratamiento de bases de datos, busca estrategias para detectar, clasificar y predecir los delitos, como lo sugieren varios autores (Sukhija et al., 2022; Suruliandi et al., 2019). Investigaciones previas en este tema abarcan variables territoriales o geoespaciales centradas en un lugar, ya sean ciudades o barrios específicos (Ordóñez et al., 2020; Aguirre et al., 2018; Aguilar & Contreras, 2018); no obstante, si bien contribuyen a entender el comportamiento delictivo en Colombia, tienen limitaciones debido a que dejan de lado criterios tan relevantes como el análisis multivariado de diferentes tipologías de delitos en diversas zonas de Colombia.

Considerando lo anterior, la investigación ha buscado responder las siguientes preguntas: ¿Cuál es la caracterización de los delitos o factores de violencia en los diferentes departamentos de Colombia? ¿Cuáles son los perfiles departamentales con mayores delitos violentos en Colombia? ¿Cómo pronosticar los perfiles de delitos en Colombia a través de inteligencia artificial? Con esto se busca establecer perfiles de violencia departamentales en los que prevalezcan delitos violentos como las amenazas, delitos sexuales, extorsión, homicidios, secuestros y terrorismo, además de la predicción basada en redes neuronales que posibiliten su pronóstico. Esto derivó en los siguientes objetivos de investigación: 1) caracterizar los delitos o factores de violencia en los diferentes departamentos de Colombia; 2) proponer los conglomerados o perfiles de delitos violentos en Colombia por medio de *clustering*, y 3) pronosticar los perfiles de delitos en Colombia mediante la inteligencia artificial.

## Marco teórico

### Factores y delitos violentos en Colombia

Los factores de violencia en Colombia han prevalecido por diversos motivos. Sánchez (2020) explica que se deben principalmente a la inequidad en el acceso a derechos básicos, así como la desigualdad en el ejercicio de la justicia, de ingresos y de acceso a oportunidades. Esto se refleja en la disparidad en calidad de vida, calidad en salud, consumo per cápita y educación en los territorios y en el contexto regional de la población. Por tanto, es importante entender las épocas y situaciones que predisponen los índices de violencia en Colombia. Un punto a tener en consideración son los delitos que implican violencia, entendida como aquellas acciones que implican comportamientos dañinos sobre una persona, que se presentan en delitos como los homicidios, la violencia sexual y los hurtos (Magalhães, 2022).

Por ende, el análisis de delitos por tipología y ubicación espacial y temporal es relevante para desplegar las operaciones policiales necesarias en el momento exacto. De ahí que se tenga conciencia de las áreas de bajo y alto impacto delictivo según su ubicación, patrones y épocas del año (Butt et al., 2021). Por ello, todos los países establecen normas que tipifican estos incidentes con valor judicial. En Colombia, los delitos se establecen en el Código Penal (Ley 599, 2000), donde prevalecen las conductas punibles y sus sanciones, con el fin de regular la sociedad y garantizar la acción de la justicia. La Policía Nacional clasifica y registra los eventos delictivos (como homicidios, lesiones, hurtos, piratería, abigeato, terrorismo, delitos sexuales, violencia intrafamiliar y amenazas) que ocurren en el territorio colombiano a través de su base de información de estadísticas delictivas, donde los desagrega en variables como tiempo, modo y lugar, con el objetivo de entender el desarrollo de cada crimen cometido (Policía Nacional de Colombia, 2022). Con esto, para efectos prácticos de la investigación, fue necesario definir los delitos violentos objeto de estudio, de acuerdo con el Código Penal colombiano (Ley 599, 2000).

Así, las amenazas se consideran delitos que atemorizan a personas, instituciones, entidades o familias con el fin de propagar terror o coaccionarlas; los delitos sexuales son actos carnales abusivos y violentos; los delitos de extorsión son un acto de opresión y de obligación a tolerar o pasar por alto hechos, palabras y órdenes con el propósito de recibir beneficios para sí mismo o terceros. Por otra parte, los homicidios son delitos que atentan contra la vida de una persona, ya sea homicidio culposo, preterintencional o por piedad; a su vez, los delitos de secuestro se refieren a actos de privación de la libertad por retener, arrebatar o sustraer a una persona, ya sea secuestro simple o extorsivo. Por último, el terrorismo procesa aquellos actos de conflicto armado en el que estén implicados ataques de violencia hacia la población.

### ***Clustering* en el sector de criminología**

El análisis de *clustering* (o agrupamiento) es una técnica multivariante enfocada en formar grupos o conglomerados de datos, con el fin de obtener etiquetas de características diferenciadoras para su comparación, con base en la cercanía o similitud de sus elementos para conocer los patrones estudiados. De acuerdo al concepto de Chang y Chen (2018), la técnica de clústeres, denominada también como análisis de segmentación o de taxonomía, aplica un método estadístico para generar grupos a través de una muestra de datos, señalando el número adecuado de conglomerados para evaluar variables de la mejor forma y establecer aquellos grupos similares que se alejan o son diferentes entre sí. Otros investigadores consideran este método como la partición de variables en subconjuntos para un agrupamiento que permita hallar la conformación organizada y brindar hipótesis en respuesta a las inquietudes surgidas al estudiar o investigar dichas variables (Jami & Gore, 2019).

En este sentido, las técnicas de *clustering* se han empleado para evaluar diversos temas. De la Hoz-Domínguez et al. (2021) utilizaron el análisis por conglomerados para identificar los grupos de 27 entidades de salud acreditadas en alta calidad, siguiendo los patrones de acreditación y variables de entrada como cuentas por cobrar, inventario, propiedad y equipo, y variables de salida como las ventas de servicios de salud y utilidad neta. Esto arrojó dos principales grupos de instituciones: las acreditadas en proceso de consolidación financiera y otras instituciones acreditadas financieramente sólidas. Fontalvo et al. (2021) también mostraron la importancia del *clustering* en análisis de diversos contextos. Igualmente, existen estudios de aplicación de *clustering* en el campo criminalístico: Sukhija et al. (2022) identificaron puntos criminales con clasificadores no supervisados y el algoritmo *K-nearest* (K-más cercano) en Gurgaon, India, teniendo en cuenta casos delictivos de los últimos cinco años en esta zona; de esta forma se reconocieron los puntos geográficos en función de la captación de delitos, con una precisión del 99 %.

Por su parte, Butt et al. (2021) dispusieron de un análisis de conglomerado jerárquico mediante un algoritmo de agrupamiento HDDSCAN, que filtra datos espaciales para encontrar las características de los grupos sin información *a priori* sobre el número de clústeres que deben ser encontrados dentro de la región —en este caso, las zonas de mayor

población, negocios y movilidad de Nueva York—. Los resultados demuestran la zona con mayor delincuencia, lo que permite empoderar a las autoridades para que desarrollen patrullajes o estrategias basadas en esta información. Otros trabajos usan estas herramientas para investigar más a profundidad a partir de los puntos calientes localizados por clústeres, con el fin de predecir sus ocurrencias en un futuro, en conjunto con análisis bidimensionales, lo que permite una mayor comprensión y actuación sobre los espacios locales según la categoría delictiva y su relación con variables como el tiempo, la ubicación, los ingresos por familia y demás (Hajela et al., 2020).

Por último, en Colombia también se han hecho estudios de análisis por grupos en diferentes ciudades. Tal es el caso expuesto por Aguirre et al. (2018), que demostraron cómo a través de la minería de datos es posible identificar barrios con mayor significancia en hurtos, así como los departamentos donde prevalecieron delitos sexuales entre los años de 2010 y 2016. Además, evidencia que estas técnicas permiten un aprovechamiento del volumen de datos tratados, dado que las técnicas tradicionales resultan ineficientes para ello; aunque sea un método complejo, la información resultante genera nuevo conocimiento para la toma de decisiones.

## **Redes neuronales y predicción en el sector de delitos en Colombia**

Las redes neuronales pertenecen al subgrupo de técnicas empleadas en inteligencia artificial y *deep learning* (aprendizaje profundo). Se trata de redes neuronales artificiales que imitan el comportamiento basado en datos históricos para predecir o pronosticar ocasiones futuras en las que se repitan y detecten tales conductas, y que adquieren experiencia a través del aprendizaje automático en la entrada de nodos, nodos ocultos y nodos de salida, que permiten arrojar un resultado concreto. Qi et al. (2019) explican que las redes neuronales son un tipo de inteligencia artificial inspeccionada, con etiquetas en las que los datos de entrada predisponen las respuestas de los modelos estudiados.

Esto ha permitido el desarrollo de nuevas investigaciones, como la detección de comportamientos maliciosos en equipos Android mediante la toma de muestra de acciones y su integración por grupos de características por cada comportamiento, para entrenar la red neuronal y finalmente identificar en la vida real aquellos patrones que predicen un enlace malicioso (Chen et al., 2020). Asimismo, otros autores utilizan las redes neuronales para detectar otros fenómenos, como la identificación del *bullying* mediante el reconocimiento de expresiones o palabras que puedan convertirse en ciberacoso, para así ayudar a la población víctima de esta problemática (Cumba-Armijos et al., 2022).

Asimismo, el uso de las redes neuronales en análisis y predicción delincriminal ha prevalecido en los últimos años, por lo cual se han incrementado los países que lo han aplicado, así como los *softwares* y artículos publicados para la mejor detección y acción sobre los hechos delictivos (Walteros et al., 2021; Khan & Wani, 2019). Es así como se ha podido establecer una red neuronal capaz de detectar intrusiones en la red con un algoritmo de red difusa y métodos de identificación lineal, con resultados exactos, para asegurar y proteger la red (Dai; 2020).



Por su parte, George et al. (2019) compilaron una red neuronal de dos clases que predice cuándo una persona habla y se expresa con sinceridad o falsedad a partir de sus gestos faciales. Primero, se entrenó la red con 62 sujetos de pruebas y se almacenó la variabilidad muscular facial para luego probar y generar pronósticos al respecto. Estos antecedentes muestran la practicidad y funcionalidad de estas técnicas innovadoras, que garantizan un tratamiento competente de los datos que se analizan dirigido a la toma de decisiones, por cuanto se pueden pronosticar comportamientos futuros en contextos delictivos.

### Metodología

Para el desarrollo de la investigación se asumió un enfoque epistemológico explicativo-cuantitativo. Inicialmente se caracterizó, organizó y tabuló la información empírica asociada a los delitos o factores de violencia en los diferentes departamentos de Colombia en una ventana de observación de 2018 a 2022. Luego se analizaron los diferentes modelos asociados con el *clustering*, y después se propuso un modelo que permitió establecer los conglomerados o perfiles relacionados con los delitos o factores de violencia para los diferentes departamentos del país. Finalmente, con los conglomerados como variables de entrada, a través de un análisis racional lógico, se estableció una estructura que permitió instaurar una red neuronal de doble capa para predecir dichos perfiles o clústeres de los factores de violencia en Colombia.

Como información primaria, se partió de las estadísticas generadas por la Policía Nacional de Colombia en su base de datos de estadísticas delictivas. Así, se utilizaron registros delictivos de amenazas, delitos sexuales, extorsión, homicidio, secuestro y terrorismo entre 2018 y septiembre de 2022. De esta forma, se consolidó la información de los diferentes delitos violentos para procesarlos, clasificados en los 32 departamentos de Colombia, como se observa en la Tabla 1.

**Tabla 1.** Delitos por departamentos entre 2018 y septiembre de 2022

|   | Departamentos/<br>Delitos | Amenazas | Delitos<br>sexuales | Extorsión | Homicidio | Secuestro | Terrorismo |
|---|---------------------------|----------|---------------------|-----------|-----------|-----------|------------|
| 1 | Amazonas                  | 259      | 441                 | 16        | 44        | 0         | 0          |
| 2 | Antioquia                 | 12 194   | 19 130              | 4988      | 8959      | 58        | 193        |
| 3 | Arauca                    | 1981     | 1014                | 443       | 568       | 40        | 160        |
| 4 | Atlántico                 | 10 991   | 6339                | 756       | 2306      | 14        | 3          |
| 5 | Bolívar                   | 8781     | 5137                | 760       | 1611      | 4         | 11         |
| 6 | Boyacá                    | 6637     | 3703                | 851       | 338       | 2         | 5          |
| 7 | Caldas                    | 2208     | 3249                | 360       | 689       | 2         | 6          |
| 8 | Caquetá                   | 2702     | 1507                | 391       | 611       | 3         | 12         |

Continúa tabla...

|    | Departamentos/<br>Delitos | Amenazas       | Delitos<br>sexuales | Extorsión     | Homicidio     | Secuestro  | Terrorismo |
|----|---------------------------|----------------|---------------------|---------------|---------------|------------|------------|
| 9  | Casanare                  | 1204           | 1306                | 483           | 281           | 6          | 7          |
| 10 | Cauca                     | 10 132         | 3626                | 1460          | 3079          | 61         | 80         |
| 11 | Cesar                     | 8536           | 3225                | 603           | 952           | 16         | 28         |
| 12 | Chocó                     | 2779           | 939                 | 562           | 994           | 17         | 39         |
| 13 | Córdoba                   | 8893           | 3591                | 582           | 1417          | 2          | 6          |
| 14 | Cundinamarca              | 53 626         | 30 823              | 5354          | 5840          | 41         | 22         |
| 15 | Guainía                   | 233            | 205                 | 41            | 16            | 0          | 0          |
| 16 | Guaviare                  | 953            | 554                 | 126           | 109           | 1          | 15         |
| 17 | Huila                     | 6010           | 4562                | 724           | 1048          | 11         | 4          |
| 18 | Guajira                   | 2976           | 1233                | 414           | 798           | 14         | 5          |
| 19 | Magdalena                 | 5950           | 3189                | 571           | 1149          | 10         | 5          |
| 20 | Meta                      | 9574           | 4025                | 1594          | 1089          | 16         | 41         |
| 21 | Nariño                    | 9416           | 3709                | 855           | 2348          | 20         | 61         |
| 22 | Norte de Santander        | 9997           | 3030                | 1271          | 2192          | 54         | 96         |
| 23 | Putumayo                  | 2815           | 1379                | 305           | 675           | 5          | 7          |
| 24 | Quindío                   | 1510           | 1676                | 242           | 770           | 0          | 7          |
| 25 | Risaralda                 | 1636           | 3560                | 432           | 872           | 6          | 1          |
| 26 | San Andrés                | 423            | 296                 | 45            | 124           | 2          | 0          |
| 27 | Santander                 | 11 880         | 7027                | 761           | 1155          | 11         | 7          |
| 28 | Sucre                     | 4113           | 1943                | 377           | 672           | 3          | 2          |
| 29 | Tolima                    | 7927           | 5401                | 900           | 1147          | 7          | 5          |
| 30 | Valle                     | 28 261         | 12 361              | 3842          | 9280          | 66         | 49         |
| 31 | Vaupés                    | 95             | 150                 | 8             | 11            | 1          | 0          |
| 32 | Vichada                   | 216            | 206                 | 64            | 70            | 0          | 1          |
|    | <b>Total</b>              | <b>234 908</b> | <b>138 536</b>      | <b>30 181</b> | <b>51 214</b> | <b>493</b> | <b>878</b> |

Fuente: Policía Nacional de Colombia (2022)

Con la información consolidada, se procedió a procesar la información mediante el programa de sistematización estadística Minitab 16, con el fin de establecer los grupos característicos de violencia por conglomerado y el dendograma de los perfiles de violencia por departamentos en Colombia para el periodo correspondiente. Seguidamente, con la información del *clustering* y los conglomerados generados, se utilizó el *software* estadístico IBM SPSS 27 para proponer una red neuronal de doble capa, mediante la cual se predijo el conjunto de delitos agrupados.

## Resultados

Para alcanzar los objetivos de esta investigación, fue necesario considerar y aplicar los dos modelos presentados a continuación, y que se relacionan entre sí: el *clustering* y las redes neuronales.

### *Clustering*

Tras el análisis con múltiples modelos asociados a las técnicas de clúster, se procedió a estandarizar las variables con el fin de asignar una medida homogénea de similitud para su conglomeración por grupos que se basara en un criterio certero. Con esto se pudieron identificar los modelos que se ajustaban a la información empírica recolectada y consolidada por la Policía Nacional de Colombia para el periodo 2018-2022, con las funciones jerárquicas de distancia euclidiana, para clasificar variables similares entre sí. Además, para el análisis de clúster, se decidió emplear también el enlace de Ward, en el que cada grupo se une al centro del clúster con el que se halle el menor número de errores posibles, es decir, de la diferencia de los cuadrados de sus desviaciones. Así se obtuvieron las diversas estadísticas, que se consolidaron en un dendrograma con los clústeres resultantes. En la Tabla 2 se puede observar el paso de amalgamación y en la Figura 1, el dendrograma de los clústeres de delitos por departamento.

**Tabla 2.** Paso de amalgamación

| Paso | Conglomerados | Nivel de semejanza | Nivel de distancia | Conglomerados incorporados |    | Nuevo conglomerado | Número de observaciones en el nuevo conglomerado |
|------|---------------|--------------------|--------------------|----------------------------|----|--------------------|--------------------------------------------------|
| 1    | 31            | 99,590             | 0,0367             | 15                         | 32 | 15                 | 2                                                |
| 2    | 30            | 99,368             | 0,0565             | 1                          | 15 | 1                  | 3                                                |
| 3    | 29            | 99,102             | 0,0803             | 1                          | 31 | 1                  | 4                                                |
| 4    | 28            | 98,497             | 0,1344             | 1                          | 26 | 1                  | 5                                                |
| 5    | 27            | 98,142             | 0,1661             | 8                          | 23 | 8                  | 2                                                |
| 6    | 26            | 97,090             | 0,2602             | 7                          | 25 | 7                  | 2                                                |
| 7    | 25            | 97,075             | 0,2615             | 17                         | 19 | 17                 | 2                                                |
| 8    | 24            | 96,956             | 0,2721             | 8                          | 28 | 8                  | 3                                                |
| 9    | 23            | 96,433             | 0,3188             | 5                          | 29 | 5                  | 2                                                |
| 10   | 22            | 96,243             | 0,3358             | 8                          | 24 | 8                  | 4                                                |
| 11   | 21            | 95,523             | 0,4003             | 8                          | 9  | 8                  | 5                                                |
| 12   | 20            | 95,111             | 0,4371             | 5                          | 13 | 5                  | 3                                                |
| 13   | 19            | 94,153             | 0,5227             | 1                          | 16 | 1                  | 6                                                |

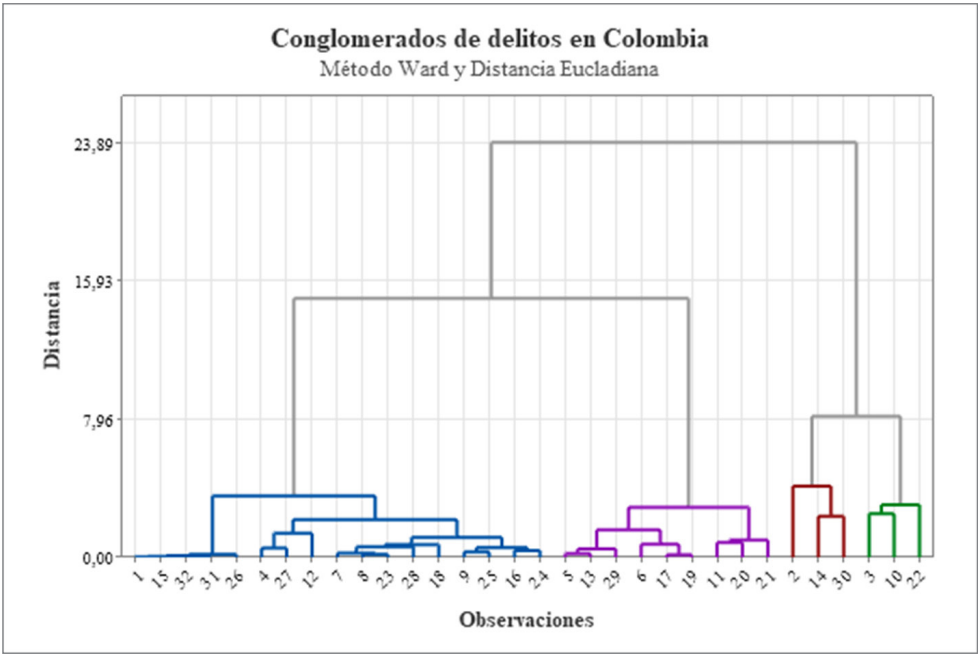
Continúa tabla...

| Paso | Conglomerados | Nivel de semejanza | Nivel de distancia | Conglomerados incorporados |    | Nuevo conglomerado | Número de observaciones en el nuevo conglomerado |
|------|---------------|--------------------|--------------------|----------------------------|----|--------------------|--------------------------------------------------|
| 14   | 18            | 93,810             | 0,5533             | 4                          | 27 | 4                  | 2                                                |
| 15   | 17            | 93,243             | 0,6040             | 7                          | 8  | 7                  | 7                                                |
| 16   | 16            | 92,687             | 0,6537             | 10                         | 22 | 10                 | 2                                                |
| 17   | 15            | 92,266             | 0,6914             | 5                          | 6  | 5                  | 4                                                |
| 18   | 14            | 92,009             | 0,7143             | 17                         | 18 | 17                 | 3                                                |
| 19   | 13            | 91,947             | 0,7199             | 11                         | 12 | 11                 | 2                                                |
| 20   | 12            | 89,657             | 0,9246             | 20                         | 21 | 20                 | 2                                                |
| 21   | 11            | 87,148             | 1,1489             | 11                         | 17 | 11                 | 5                                                |
| 22   | 10            | 85,776             | 1,2715             | 4                          | 5  | 4                  | 6                                                |
| 23   | 9             | 77,550             | 2,0069             | 4                          | 11 | 4                  | 11                                               |
| 24   | 8             | 73,808             | 2,3414             | 1                          | 7  | 1                  | 13                                               |
| 25   | 7             | 73,194             | 2,3963             | 4                          | 20 | 4                  | 13                                               |
| 26   | 6             | 68,001             | 2,8605             | 3                          | 10 | 3                  | 3                                                |
| 27   | 5             | 57,544             | 3,7953             | 2                          | 30 | 2                  | 2                                                |
| 28   | 4             | 35,062             | 5,8051             | 2                          | 14 | 2                  | 3                                                |
| 29   | 3             | 11,718             | 7,8918             | 1                          | 4  | 1                  | 26                                               |
| 30   | 2             | -30,135            | 11,6332            | 2                          | 3  | 2                  | 6                                                |
| 31   | 1             | -207,934           | 27,5273            | 1                          | 2  | 1                  | 32                                               |

Fuente: Elaboración propia con Minitab 16.

Así pues, teniendo en cuenta las corridas junto con los niveles de distancia y de semejanza entre variables, se dividieron los departamentos por los grupos característicos más cercanos entre sí, como se observan en el dendograma (Figura 1), que estableció los cuatro conglomerados o clústeres de los diferentes delitos clasificados en las diferentes regiones de Colombia. De este modo se pudieron reconocer los perfiles y sectores con mayor o menor impacto de delitos de violencia en el país.

Luego, mediante el dendograma, se pudieron reconocer cuatro clústeres. El primer clúster hace referencia a los departamentos de Amazonas, Guainía, Vichada, Vaupés, San Andrés, Guaviare, Caldas, Risaralda, Caquetá, Putumayo, Sucre, Quindío y Casanare. En el segundo clúster se encuentran los departamentos de Antioquia, Valle y Cundinamarca. El tercer clúster agrupa los departamentos de Arauca, Cauca y Norte de Santander. Por último, en el cuarto clúster se encuentran los departamentos de Atlántico, Santander,



**Figura 1.** Dendrograma de los clústeres de violencia por departamento.

Fuente: Elaboración propia con Minitab 16

Bolívar, Tolima, Córdoba, Boyacá, Cesar, Chocó, Huila, Magdalena, Guajira, Meta y Nariño. De esta manera se clasificaron los diferentes departamentos de Colombia en los grupos jerárquicos. Además, en la Tabla 3, se evalúa el número de departamentos correspondiente a cada conglomerado, sus distancias respecto al centro y su pertenencia a las características particulares encontradas en los clústeres que aluden a los criterios seleccionados para corresponder a cada grupo.

**Tabla 3.** Partición final por clúster o perfil de delitos

| Clúster   | Número de observaciones | Dentro de la suma de cuadrados del conglomerado | Distancia promedio desde el centroide | Distancia máxima desde el centroide |
|-----------|-------------------------|-------------------------------------------------|---------------------------------------|-------------------------------------|
| Clúster 1 | 13                      | 1,3628                                          | 0,31678                               | 0,50325                             |
| Clúster 2 | 3                       | 23,9197                                         | 2,75946                               | 3,33842                             |
| Clúster 3 | 3                       | 3,7537                                          | 1,03429                               | 1,53623                             |
| Clúster 4 | 13                      | 6,2366                                          | 0,64984                               | 1,17533                             |

Fuente: Elaboración propia con Minitab 16

En la Tabla 3 se identifican también los clústeres según su incidencia en hechos delictivos. El clúster con mayor incidencia es el 2, seguido por el clúster 3, luego el clúster 4 y finalmente el clúster de menor incidencia es el 1, a pesar de ser uno de los grupos con mayor número de departamentos. Estos resultados muestran que el clúster 1 contiene el conjunto de departamentos con menores niveles de violencia y ocurrencia de delitos en la ventana de tiempo de evaluación de esta investigación. Por su parte, el clúster 2, aun cuando lo componen solo tres departamentos, reúne el mayor número de delitos de violencia cometidos en Colombia en el periodo evaluado. En consecuencia, en la Tabla 4 se señalan los centroides de los conglomerados asociados con los delitos de amenazas, delitos sexuales, extorsión, homicidio, secuestro y terrorismo.

**Tabla 4.** Centroides de grupo de los clústeres de delitos por departamento

| Variable         | Clúster 1 | Clúster 2 | Clúster 3 | Clúster 4 | Centroide principal |
|------------------|-----------|-----------|-----------|-----------|---------------------|
| Amenazas         | -0,581995 | 2,35815   | 0,00286   | 0,037146  | -0,0000000          |
| Delitos sexuales | -0,497539 | 2,67149   | -0,28801  | -0,052496 | 0,0000000           |
| Extorsión        | -0,549340 | 2,88433   | 0,08752   | -0,136472 | 0,0000000           |
| Homicidio        | -0,535917 | 2,82245   | 0,15193   | -0,150476 | 0,0000000           |
| Secuestro        | -0,664814 | 1,99784   | 1,82964   | -0,218451 | -0,0000000          |
| Terrorismo       | -0,498491 | 1,31398   | 1,83468   | -0,228123 | 0,0000000           |

Fuente: Elaboración propia con Minitab 16

En la Tabla 4, se definieron los grupos por grado de abordaje de delitos. Al respecto, el clúster 2 confirma que la mayor repercusión de violencia e inseguridad prevaleció por cada tipo de delito. Al mismo tiempo, el clúster 3 señaló niveles altos de los diferentes delitos, a pesar de que los delitos sexuales presentan menores repercusiones sobre estos departamentos. Por otra parte, el clúster 4 mostró niveles negativos para los diferentes delitos, lo cual contrasta con los clústeres 2 y 3. No obstante, en este clúster las amenazas sobresalen sobre los demás delitos, aunque, a diferencia de los otros clústeres, su nivel es muy bajo. En suma, los clústeres 2 y 3 tienden a ser departamentos altamente delictivos, y el menor puntaje delictivo entre todos se evidenció en los departamentos del clúster 1.

En la Figura 1, el dendrograma de los cuatro clústeres de violencia por departamento en Colombia tiende a dividirse en dos grandes grupos que aglomeran la mayoría de los departamentos, y que corresponden a los clústeres 1 y 4. Sin embargo, al revisar la Tabla 5, el clúster 2 (Antioquia, Valle y Cundinamarca) muestra un conglomerado que estadísticamente resalta la mayor distancia de centroide, es decir, posee los departamentos con

mayor índice de violencia en Colombia. Así mismo, al revisar el clúster 3 (Arauca, Cauca y Norte de Santander), se aprecia un centroide que presenta el segundo valor más alto de los cuatro conglomerados, de lo cual se concluye que es el segundo grupo con mayor impacto en niveles de delitos cometidos en Colombia. En consecuencia, la relación del clúster 2 con los otros tres demostró una cercanía con los departamentos del clúster 3. Ello significa que las distancias entre los grupos 4 y 1 adquieren características similares, que pueden incrementarse o reducirse en un futuro.

**Tabla 5.** Las distancias entre los centroides de clústeres

|           | Clúster 1 | Clúster 2 | Clúster 3 | Clúster 4 |
|-----------|-----------|-----------|-----------|-----------|
| Clúster 1 | 0,00000   | 7,22016   | 3,59593   | 1,08294   |
| Clúster 2 | 7,22016   | 0,00000   | 5,43685   | 6,16934   |
| Clúster 3 | 3,59593   | 5,43685   | 0,00000   | 2,94077   |
| Clúster 4 | 1,08294   | 6,16934   | 2,94077   | 0,00000   |

Fuente: Elaboración propia con Minitab 16

### Redes neuronales

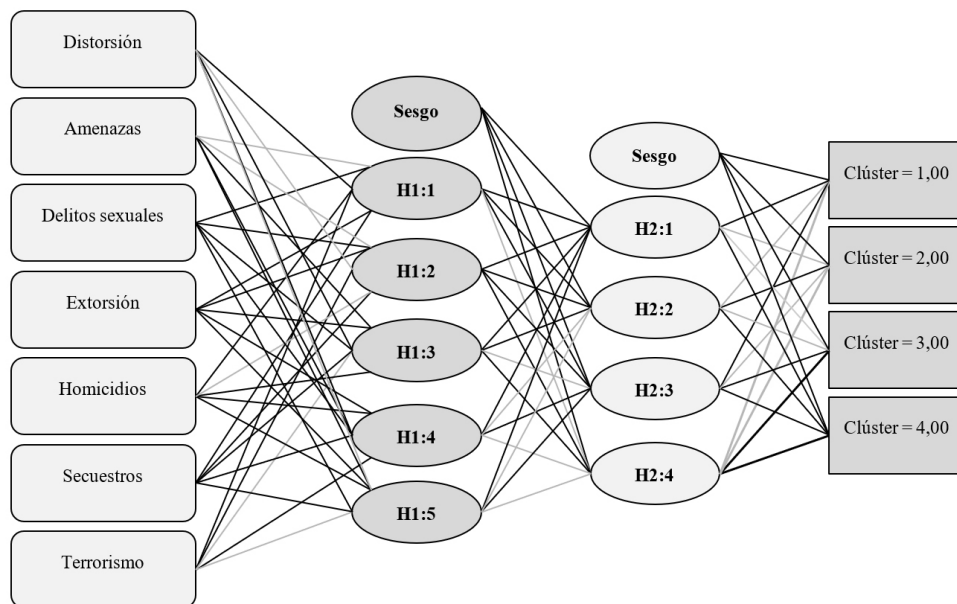
Con el fin de establecer un modelo para pronosticar los perfiles de violencia, en este apartado se desarrolló el procedimiento para implementar una red neuronal. Para iniciar, se precisaron los nodos de entrada y salida. De este modo, se procedió a establecer una estructura de inteligencia artificial, concebida mediante una red neuronal de doble capa. Para establecerla, se tomaron como punto de partida los cuatro clústeres, señalados en la Tabla 5. A su vez, como variables independientes se procesaron los clústeres identificados en el apartado anterior, además, como covariables se tuvieron en cuenta los diferentes delitos analizados. Finalmente, la estructura del modelo establecido trabajó como capa de entrada un modelo de tangente hiperbólica para ingresar los patrones que se comunican a la red neuronal, mientras que las capas ocultas dieron como resultado la capa de salida, empleando para esta el modelo sigmoide. Como función de error se utilizó la suma de cuadrados.

Enseguida, al ingresar los datos en el programa y procesarlos, se entrenó la red neuronal con los datos de muestra y se completó la corrida de data al 100 % de valores válidos sin exclusión alguna para mayor precisión de respuesta. En consecuencia, los valores de la Tabla 6 generaron la estructura que se muestra en la Figura 2, correspondiente a la red de doble capa en la red neuronal. Allí se evidenció el comportamiento de los delitos con respecto a los clústeres expuestos, con el fin de pronosticar la pertenencia a un determinado conglomerado.

**Tabla 6.** Resumen del procesamiento de los casos

|                |               | N.º       | Porcentaje |
|----------------|---------------|-----------|------------|
| <b>Muestra</b> | Entrenamiento | 13        | 40,6 %     |
|                | Prueba        | 15        | 46,9 %     |
|                | Reserva       | 4         | 12,5 %     |
| Válidos        |               | 32        | 100,0 %    |
| Excluidos      |               | 0         |            |
| <b>Total</b>   |               | <b>32</b> |            |

Fuente: Elaboración propia con Minitab 16



**Figura 2.** Estructura de red neuronal doble capa para pronosticar los delitos a un conglomerado determinado.

Fuente: Elaboración propia mediante IBM SPSS 27

Por otra parte, en la Tabla 7 se observó una alta capacidad de clasificación del modelo propuesto para la partición de las 32 muestras, para su entrenamiento, prueba y reserva, en razón a los departamentos de Colombia y su respectivo porcentaje pronosticado por cada clúster encontrado. Esto probó una capacidad promediada de clasificación para el universo estudiado del 97,7 %, un resultado bueno para el modelo propuesto.



**Tabla 7.** Capacidad de clasificación del modelo RNA de doble capa

| Muestra                 | Observado            | Pronosticado |        |        |        | Porcentaje<br>correcto |
|-------------------------|----------------------|--------------|--------|--------|--------|------------------------|
|                         |                      | 1,00         | 2,00   | 3,00   | 4,00   |                        |
| Entre-<br>namiento      | 1,00                 | 3            | 0      | 0      | 0      | 100,0 %                |
|                         | 2,00                 | 0            | 1      | 0      | 0      | 100,0 %                |
|                         | 3,00                 | 0            | 0      | 1      | 0      | 100,0 %                |
|                         | 4,00                 | 0            | 0      | 0      | 8      | 100,0 %                |
|                         | Porcentaje<br>global | 23,1 %       | 7,7 %  | 7,7 %  | 61,5 % | 100 %                  |
| Prueba                  | 1,00                 | 6            | 0      | 0      | 1      | 85,7 %                 |
|                         | 2,00                 | 0            | 1      | 0      | 0      | 100,0 %                |
|                         | 3,00                 | 0            | 0      | 2      | 0      | 100,0 %                |
|                         | 4,00                 | 0            | 0      | 0      | 5      | 100,0 %                |
|                         | Porcentaje<br>global | 40,0 %       | 6,7 %  | 13,3 % | 40,0 % | 93,3 %                 |
| Reserva                 | 1,00                 | 3            | 0      | 0      | 0      | 100,0 %                |
|                         | 2,00                 | 0            | 1      | 0      | 0      | 100,0 %                |
|                         | 3,00                 | 0            | 1      | 0      | 0      | 0,0 %                  |
|                         | 4,00                 | 0            | 0      | 0      | 2      | 0,0 %                  |
|                         | Porcentaje<br>global | 75,0 %       | 25,0 % | 0,0 %  | 0,0 %  | 100,0 %                |
| Porcentaje<br>ponderado |                      |              |        |        |        | 97,7 %                 |

Fuente: Elaboración propia mediante IBM SPSS 27

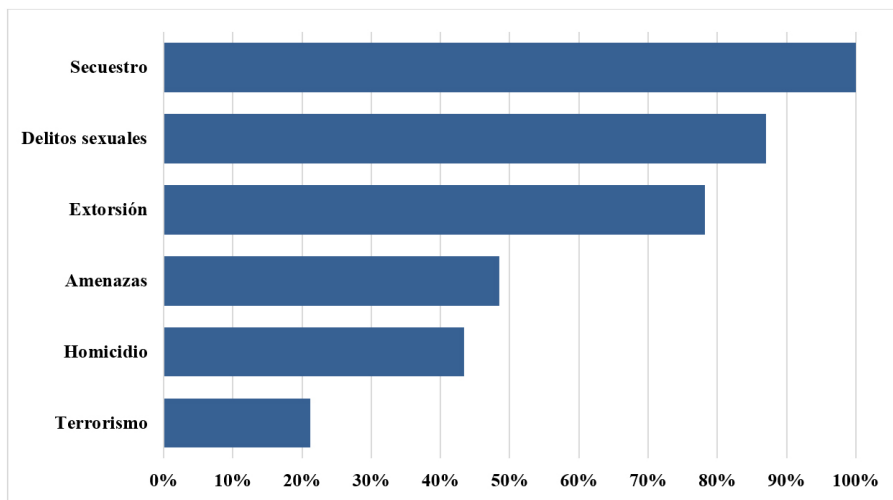
Así mismo, en la Tabla 8 se evidencia el nivel de importancia relativa de los diferentes delitos analizados ante la estructura de doble capa relacionada con la red neuronal propuesta. Se observó que el delito de mayor importancia relativa para la generación del comportamiento y los patrones es el secuestro, mientras que el de menor incidencia, relevancia e impacto dentro del modelo es el terrorismo. Esto es analizado desde la información estudiada y los resultados cuantitativos del modelo.

**Tabla 8.** Importancia de las variables independientes

| Delitos          | Importancia | Importancia normalizada |
|------------------|-------------|-------------------------|
| Amenazas         | ,128        | 48,5 %                  |
| Delitos sexuales | ,230        | 87,0 %                  |
| Extorsión        | ,207        | 78,2 %                  |
| Homicidio        | ,115        | 43,4 %                  |
| Secuestro        | ,264        | 100,0 %                 |
| Terrorismo       | ,056        | 21,2 %                  |

Fuente: Elaboración propia mediante IBM SPSS 27

Del mismo modo, para mayor comprensión, se explicitan los datos de la Tabla 8 de manera gráfica en la Figura 3, donde se visualizan los delitos desde un alto hasta un bajo nivel de significancia para el modelo de doble capa de red neuronal, de manera que se pudieron pronosticar los clústeres de los delitos en los diferentes departamentos de Colombia con las condiciones de importancia expuestas.



**Figura 3.** Importancia normalizada de los delitos.

Fuente: Elaboración propia mediante IBM SPSS 27

Como hallazgo de esta investigación, se encontró que en los departamentos de Cundinamarca, Antioquía y Valle se generan los mayores niveles de violencia en Colombia. Del mismo modo, el segundo clúster de mayor nivel de violencia es el conformado por los departamentos de Arauca, Cauca y Norte de Santander, aunque muestran una menor implicación de ocurrencia de estos delitos. Como hallazgo asociado a la estructura de red

neuronal, se evidenció que la estructura propuesta muestra una capacidad de pronosticar a qué conglomerado podría pertenecer o pasar un departamento a futuro con una probabilidad del 97,7%.

De igual manera, las variables estandarizadas para la red neuronal mostraron que los delitos de mayor importancia para el modelo propuesto en Colombia, en su orden, fueron el secuestro, los delitos sexuales, la extorsión, las amenazas, los homicidios y el terrorismo, como se observa en la Tabla 8 y la Figura 3. En consecuencia, por los altos índices de violencia y los múltiples delitos que se cometen en Colombia, este análisis, que integró el cálculo multivariable y la red neuronal, se constituye en un aporte significativo para la Policía Nacional y demás entidades relacionadas, con el fin de seguir investigando sobre el tema, para optimizar la toma de decisiones relacionadas con la disminución de los delitos analizados en Colombia.

## Discusión

Otras investigaciones ya han demostrado la pertinencia de estas técnicas en otros contextos de incidencia pública (Fontalvo et al., 2018), lo que ratifica la importancia y utilidad de los resultados de esta investigación. Otros estudios han aplicado las técnicas de conglomerado en contextos similares. Un ejemplo claro es la investigación de Das et al. (2021), en la que se agruparon clústeres de informes delictivos para un tratamiento eficaz. La adaptabilidad de estos grupos fue posible al emplear un conjunto de redes neuronales, lo que hizo factible la actualización inmediata al ingresar otros informes delictivos con niveles de rendimiento altos, así como en este estudio se priorizó el empleo de datos actualizados.

Además, Deepak et al. (2021) evidencian que este tipo de métodos alcanzan métricas de respuesta precisas en comparación con otras técnicas de evaluación, al clasificar los tipos de delitos mediante procesos de redes neuronales. Por otra parte, estudios han utilizado los modelos de redes neuronales artificiales para predecir los tipos de robos considerando factores ambientales; sin embargo, explican que esto es posible solo si se integran las características y variables relacionadas con ese tipo de delito. Igualmente, resaltan la importancia de incursionar en áreas criminalísticas para proporcionar diversidad a los modelos de redes neuronales (Kwon et al., 2021), tal como se hizo en esta investigación.

En cuanto al método de *clustering*, la mayoría de investigaciones aplican una cantidad de dos clústeres (Azis et al., 2021). Si bien esto permite focalizar el análisis en contextos con mayores y menores delitos, deja de lado una jerarquización más desagregada de regiones que podrían clasificarse y desglosarse según las cualidades propias de los delitos, de manera que permita a las autoridades comprender a profundidad el comportamiento de cada uno de los grupos e incluso comprender el posible desarrollo de capacidades de violencia diferentes. En este estudio, si bien dos clústeres agruparon la mayoría de departamentos, los otros dos fueron relevantes en materia de cantidad de delitos, lo que permite priorizar las acciones de las entidades encargadas de la seguridad en el país.

Un punto a destacar de los resultados obtenidos en este estudio es la precisión de la red neuronal propuesta para la proyección de los delitos. Por tanto, la calidad y precisión de la información ha sido significativa, ya que permitió entrenar la red neuronal de forma efectiva. En este sentido, dependiendo de la calidad de los datos utilizados, el sistema de red neuronal propuesto será más o menos preciso. En otras investigaciones se ha demostrado que el empleo de redes neuronales permite alcanzar una buena capacidad de predicción. Tal es el caso de Incio et al. (2021), que utilizaron dos algoritmos y obtuvieron como resultado un porcentaje de predicción del 70 % para el primero y de 86 % para el segundo en una red neuronal de doble capa. Por su parte, Bravo et al. (2021) proponen una red neuronal con capacidad de predicción del 97,3 % al entrenar y aumentar el análisis de los datos. En comparación con dichos estudios, esta investigación demostró que la proyección de los delitos de los cuatro clústeres propuestos tiene una capacidad de predicción del 97,7 % sobre la pertenencia de un departamento a uno de los cuatro conglomerados de delitos.

Por último, esta alta capacidad de identificación y pronóstico de la red neuronal propuesta puede permitir que las autoridades encargadas de la seguridad y bienestar social en Colombia comprendan el comportamiento de los delitos por departamentos en estos últimos cuatro años, para predecir a su vez el panorama futuro de Colombia al respecto. De esta forma se pueden identificar los departamentos en que se deben priorizar medidas de seguridad y de preparación de la población residente, así como tomar decisiones y establecer políticas de Estado para reducir dichos factores de violencia.

## Conclusión

Como aporte teórico y metodológico, esta investigación articuló los delitos y factores de violencia en Colombia con la aplicación del método de *clustering*, lo cual permitió identificar los clústeres de violencia en el periodo 2018-2022. Se brinda así un análisis que permitió caracterizar los delitos violentos mediante el enlace de grupos de factores de violencia por departamentos, para finalmente proyectar la pertenencia de este conjunto de delitos a un conglomerado específico, lo que permitió identificar el comportamiento de estos factores de violencia por grupos de departamentos.

Como aporte práctico, las entidades del Estado responsables de la toma de decisiones en seguridad y de prevenir la ocurrencia de los delitos, recibieron los criterios cuantitativos soportados en el cálculo multivariado y las redes neuronales, lo que permite identificar los puntos geográficos críticos de Colombia para establecer políticas públicas que prioricen esos grupos de departamentos con mayores índices de violencia.

De igual manera se recomienda hacer análisis similares de las causas de los factores de violencia a nivel nacional y departamental para contrastar estos resultados, en pro de la resolución de esta problemática y su prevalencia en los próximos años. Asimismo, se sugiere identificar en estudios futuros otros conglomerados y tipologías diferentes a los delitos estudiados acá, con base en otros modelos de aprendizaje automático o *machine learning*,

de modo que se analicen desde otras perspectivas los delitos y su eventual predicción (Shermila et al., 2018). De esta forma, se puede avanzar en el desarrollo de estructuras de predicción con mayor capacidad de precisión.

Esta investigación es de utilidad a nivel regional, nacional e internacional, dado que la metodología y estructura propuesta en esta investigación se puede replicar en diferentes contextos, como barrios, municipios, departamentos y otros países, para clasificar contextos específicos y hacer pronósticos en pro de la toma de decisiones para reducir los diferentes tipos de delitos.

### Agradecimientos

Los autores desean agradecer a la Universidad de Cartagena y la Universidad del Tolima por su apoyo en la realización de este artículo.

### Declaración de divulgación

Los autores declaran que no existe ningún potencial conflicto de interés relacionado con el artículo.

### Financiamiento

Los autores no declaran fuente de financiamiento para la realización de este artículo.

### Sobre los autores

**Tomás José Fontalvo Herrera** es doctor en ciencias sociales, Universidad del Zulia, Venezuela, y magíster en administración, Universidad Nacional de Colombia. Profesor de tiempo completo de la Universidad de Cartagena. Es conferencista, consultor e investigador en el campo de gestión de calidad, control y cálculo multivariante, seis sigmas, eficiencia, productividad, gestión y *machine learning*.

<https://orcid.org/0000-0003-4642-9251>

Contacto: [tfontalvoh@unicartagena.edu.co](mailto:tfontalvoh@unicartagena.edu.co)

**María Alejandra Vega Hernández** es administradora industrial con énfasis en producción de la Universidad de Cartagena e investigadora junior de calidad integral y cálculo multivariante. Hace parte del Grupo de Investigación de Productividad Organizacional de la Universidad de Cartagena.

<https://orcid.org/0000-0001-5898-7440>

Contacto: [mvegah1@unicartagena.edu.co](mailto:mvegah1@unicartagena.edu.co)

**Fabio Mejía Zambrano** es candidato doctor en administración; magíster en mercadeo agroindustrial, y magíster en docencia universitaria. Profesor asociado y miembro del Grupo de Investigación ECO de la Universidad del Tolima. Investigador en planeación prospectiva y sustentabilidad en educación superior. Ha sido director de plan de desarrollo y vicerrector académico en universidades públicas, y consultor en gerencia.

<https://orcid.org/0000-0002-6689-3345> - Contacto: [fmejiaz@ut.edu.co](mailto:fmejiaz@ut.edu.co)

## Referencias

- Aguilar, A., & Contreras, M. (2018). *Caracterización de los delitos en Cartagena mediante la aplicación de minería de datos* [tesis doctoral, Universidad Tecnológica de Bolívar]. <https://biblioteca.utb.edu.co/notas/tesis/0074619.pdf>
- Aguirre Yacup, N., Walteros Alcázar, M., & Noguera Pérez, Y. (2018). Identificación de patrones delictivos en Colombia durante el periodo 2010-2016 mediante el uso de técnicas de minería de datos. *Encuentro Internacional de Educación en Ingeniería Acofi 2018*. <https://acofipapers.org/index.php/eiei/article/view/435>
- Azis, N., Ali, Y., Subekti, R., Junianto, P., Diner, L., Suhendra, S., Aisyah, S., & Windarto, A. P. (2021, febrero). Mapping study using the unsupervised learning clustering approach. *IOP Conference Series: Materials Science and Engineering*, 1088(1), p. 012005. <https://doi.org/10.1088/1757-899X/1088/1/012005>
- Bravo, M., Artega, H., Tabares, K., Padilla, I., & Orozco, S. (2021). Cervical cancer classification using convolutional neural networks, transfer learning and data augmentation. *Revista EIA*, 18(35), 100-111. <https://doi.org/10.24050/reia.v18i35.1462>
- Butt, U., Letchmunan, S., Hassan, F., Ali, M., Baqir, A., Koh, T., & Sherazi, H. (2021). Spatio-temporal crime predictions by leveraging artificial intelligence for citizens security in smart cities. *IEEE Access*, 9, 47516-47529. <https://doi.org/10.1109/ACCESS.2021.3068306>
- Chang, D. F., & Chen, C. C. (2018). Cluster analysis for student performance in PISA2015 among OECD economies. *ICIC Express Letters, Part B, Applications: An International Journal of Research and Surveys*, 9(11), 1139-1146. <https://doi.org/10.24507/iciclb.09.11.1139>
- Chen, T., Qiang, Y., Tang, C., Wang, Z., Luo, P., & Peng, L. (2020). Malicious behaviour identification for Android based on an RBF neural network. *International Journal of Security and Networks*, 15(3), 148-154. <http://dx.doi.org/10.1504/IJSN.2020.109706>
- Cumba-Armijos, P., Riofrío-Luzcando, D., Rodríguez-Arboleda, V., & Carrión-Jumbo, J. (2022). Detecting cyberbullying in Spanish texts through deep learning techniques. *International Journal of Data Mining, Modelling and Management*, 14(3), 234-247.
- Dai, Z. (2020). Research on network intrusion detection security based on improved extreme learning algorithms and neural network algorithms. *International Journal of Biometrics*, 12(1), 56-66.
- Das, P., Das, A., Nayak, J., Pelusi, D., & Ding, W. (2021). Group incremental adaptive clustering based on neural network and rough set theory for crime report categorization. *Neurocomputing*, 459, 465-480. <https://doi.org/10.1016/j.neucom.2019.10.109>
- Datosmacro. (2021). *Índice de Gini 2021* [página web]. <https://bit.ly/41UJ297>
- De la Hoz-Domínguez, E., Fontalvo, T., & Fontalvo, O. (2021). Methodology of classification, forecast and prediction of healthcare providers accredited in high quality in Colombia. *International Journal of Productivity and Quality Management*, 33(1), 1-20. <https://doi.org/10.1504/IJPQM.2021.115290>
- Deepak, G., Rooban, S., & Santhanavijayan, A. (2021). A knowledge centric hybridized approach for crime classification incorporating deep bi-LSTM neural network. *Multimedia Tools and Applications*, 80(18), 28061-28085. <https://doi.org/10.1007/s11042-021-11050-4>
- Departamento Administrativo Nacional de Estadística (DANE). (2021, 3 de noviembre). *Encuesta de Convivencia y Seguridad Ciudadana (ECSC). Victimización 2020 y percepciones 2021* [rueda de prensa, video]. YouTube. <https://bit.ly/45hMNZn>
- Duffield, M. (2020). Seguridad humana: vincular desarrollo y seguridad en una era de terror. *Relaciones Internacionales*, 43, 11-32. <https://doi.org/kbst>
- Fontalvo, T., De la Hoz-Domínguez, E., & Fontalvo, O. (2021). Assessing and forecasting method of financial efficiency in a free industrial economic zone. *International Journal of Productivity and Quality Management*, 33(2), 253-270. <https://doi.org/10.1504/IJPQM.2021.115694>

- Fontalvo, T., De la Hoz, E., & Mendoza, A.. (2018). Aplicación de minería de datos para la clasificación de programas universitarios de Ingeniería Industrial acreditados en alta calidad en Colombia. *Información Tecnológica*, 29(3), 89-96. <http://dx.doi.org/10.4067/S0718-07642018000300089>
- George, S., Pai, M., Pai, R., & Praharaj, S. (2019). Visual cues-based deception detection using two-class neural network. *International Journal of Computational Vision and Robotics*, 9(2), 132-151. <https://doi.org/10.1504/IJCVR.2019.098796>
- Gratius, S., & Rodríguez Pinzón, E. (2021). Entre seguridad humana y estatal: ¿ofrece la Unión Europea una respuesta coherente a los desafíos de seguridad en Centroamérica, Colombia y Venezuela? *Colombia Internacional*, 105, 117-145. <https://doi.org/10.7440/colombiaint105.2021.05>
- Hajela, G., Chawla, M., & Rasool, A. (2020). A clustering based hotspot identification approach for crime prediction. *Procedia Computer Science*, 167, 1462-1470. <https://doi.org/10.1016/j.procs.2020.03.357>
- Incio Flores, F., Capuñay Sánchez, D., Estela Urbina, R., Delgado Soto, J., & Vergara Medrano, S.. (2021). Diseño e implementación de una red neuronal artificial para predecir el rendimiento académico en estudiantes de Ingeniería Civil de la UNIFSLB. *Veritas et Scientia*, 10(1), 107-117. <https://doi.org/10.47796/ves.v10i1.464>
- Jami, H., & Gore, S. (2019). Use of data envelopment analysis for clustering of decision making units. *Italian Journal of Pure and Applied Mathematics*, 42, 575-587.
- Kahhat, F. (2019). *Seguridad internacional: una introducción crítica*. Fondo Editorial de la PUCP.
- Khan, A. I., & Wani, M. A. (2019). A common convolutional neural network model to classify plain, rolled and latent fingerprints. *International Journal of Biometrics*, 11(3), 257-273. <https://doi.org/10.1504/IJBM.2019.100843>
- Kwon, E., Jung, S., & Lee, J. (2021). Artificial neural network model development to predict theft types in consideration of environmental factors. *ISPRS International Journal of Geo-Information*, 10(2), 99. <https://doi.org/10.3390/ijgi10020099>
- Ley 599. (2000, 24 de julio). *Por la cual se expide el Código Penal*. Congreso de la República de Colombia. Diario Oficial 44097. Imprenta Nacional. <https://bit.ly/3S9XjLn>
- Magalhães, C. (2022). El concepto de violencia en psicología social: una revisión de la literatura. *Logos & Culturas*, 2(1), 58-78.
- Niño, C. (2020). Seguridad ciudadana en América Latina: gobernanzas criminales y dimensiones de la violencia. En F. Sánchez & N. Liendo (Eds.), *Manual de ciencia política y relaciones internacionales* (pp 207-230). Universidad Sergio Arboleda.. <https://doi.org/10.22518/book/9789585511972/ch08>
- Ordóñez, H., Cobos, C., & Bucheli, V. (2020). Modelo de machine learning para la predicción de las tendencias de hurto en Colombia. *Revista Ibérica de Sistemas e Tecnologías de Informação*, E29(5), 494-506. <https://bit.ly/3MsGcCP>
- Organización para la Cooperación y el Desarrollo Económicos (OECD). (2021). *How's life? Well-Being* [estadísticas]. OECD Stats. <https://stats.oecd.org/#>
- Palacio Chaves, J. E. (2020). *Influencia de la criminalidad sobre el crecimiento económico y la producción agrícola en Colombia durante 2010-2018*. <https://repository.unab.edu.co/handle/20.500.12749/12193>
- Policía Nacional de Colombia. (2022). *Estadística delictiva* [estadísticas]. <https://bit.ly/3S1efUL>
- Qi, X., Chen, G., Li, Y., Cheng, X., & Li, C. (2019). Applying neural-network-based machine learning to additive manufacturing: Current applications, challenges, and future perspectives. *Engineering*, 5(4), 721-729. <https://doi.org/10.1016/j.eng.2019.04.012>
- Rodríguez Pinzón, E. M., & Rodrigues, T. (2020). “Mano dura” y democracia en América Latina: seguridad pública, violencia y estado de derecho. *América Latina Hoy*, 84. <https://doi.org/10.14201/alh.21156>

- Sánchez Pirazán, M. C. (2020). *Factores determinantes de la violencia en Colombia y posibles vías de mejora* [tesis de pregrado, Universidad de los Andes]. <http://hdl.handle.net/1992/44477>
- Shermila, A. M., Bellarmine, A. B., & Santiago, N. (2018, May). Crime data analysis and prediction of perpetrator identity using machine learning approach. En *2nd International Conference on Trends in Electronics and Informatics (ICOEI)* (pp. 107-114). IEEE. <https://doi.org/10.1109/ICOEI.2018.8553904>
- Sukhija, K., Singh, S. N., Kumar, M., & Mehrotra, D. (2022). Spatial and temporal trends reveal hotspot identification of crimes using machine learning approach. *International Journal of Computational Science and Engineering*, 25(2), 174-185. <https://doi.org/10.1504/IJCSE.2022.122214>
- Suruliandi, A., David, H. B., & Raja, S. P. (2019). Preventing crimes ahead of time by predicting crime propensity in released prisoners using data mining techniques. *International Journal of Applied Decision Sciences*, 12(3), 307-336. <https://doi.org/10.1504/IJADS.2019.100433>
- Valencia Martínez, J. J. (2018). *Factores sociales que han incidido negativamente en la seguridad ciudadana de Santiago de Cali durante los últimos 20 años* [tesis, Especialización en Administración de Seguridad, Universidad Militar Nueva Granada, Bogotá]. RI UMNG. <http://hdl.handle.net/10654/21132>
- Walteros, M., Aguirre, N., Castillo, S., & Caicedo, P. (2021). General crime from the data mining point of view. A systematic literature review. *International Journal of Business Intelligence and Data Mining*, 19(3), 371-393. <https://doi.org/10.1504/IJBIDM.2021.118186>