Journal of Namibian Studies, 40 (2024): 294-313 ISSN: 2197-5523 (online)

Method To Assess, Classify And Project Theft Clusters In Colombia By Means Of Clustering And Neural Networks

Orianna Fontalvo[1], Tomás Fontalvo[2], Ana Banquez[3]

[1]Faculty of Industrial and Information Engineering, Politecnico di Milano, Italia, E-mail: oriannamireya.fontalvo@mail.polimi.it;
[2]Faculty of Economic Sciences, University of Cartagena, Colombia. E-mail: tfontalvoh@unicartagena.edu.co;
[3]Faculty of Economic Sciences, University of Cartagena, Colombia. E-mail: abanquezm@unicartagena.edu.co

ABSTRACT

The purpose of this research is to establish theft clusters in Colombia in order to assess, classify and project the profiles of violence in each region by means of clustering and neural networks. The theoretical framework is based on the articulation of artificial intelligence, neural networks, data analytics, clustering and theft in Colombia. In the methodology, a cross-sectional study was conducted based on a quantitative analysis, starting from historical data generated by the National Police of Colombia, on thefts that occurred in the 32 departments of Colombia in the year 2021. As a result, three theft clusters were established, which were used to identify the region's most and least affected by this problem. Subsequently, a double-layer neural network was proposed that allowed forecasting with an accuracy level of 96.97% the belonging of each context to a specific theft cluster that will exist in Colombia in the future.

Keywords: Theft, Colombia, data analytics, neural networks, security, artificial intelligence.

INTRODUCTION

Education This Theft is a crime that has been the subject of interest and debate in the field of criminology and criminal justice (Flórez, 2021). Throughout the history of human societies, the need to punish those who commit criminal acts has been a constant2, and several philosophical reasons have been cited to justify the function of punishment, ranging from expiation, reparation and social utility to crime prevention and resocialization of the offender (Roca y Trespaderne, 2021). However, the effectiveness of measures such as prison in reducing crime has come under scrutiny (Guevara y Jiménez, 2020), and empirical evidence suggests that the most important factor in deterring crime is the certainty of being caught rather than the severity or length of the penalty.

According to Attorney General's Office, in Bogota, theft accounted for 81% of complaints, while, in Medellin, it constituted 68%. Meanwhile, in cities such as Villavicencio, Cali, Manizales, Cartagena, Pereira, Barranquilla, Santa Marta and Cúcuta, theft comprised between 60% and 65% of the complaints. In Florencia, Neiva, Bucaramanga and Soacha, the figures ranged from 65% to 75%. Associated with the characteristics of the crimes, the authors Efendi et al. (2023) in their research have shown the importance of making investments and collective financing to ensure security and reduction of potential crimes in Indonesia, considering a series of challenges that respond to the different cultural and social contexts of the country or contexts that present different characteristics and patterns. Now, studies have addressed the problem of theft supported by robust data analytics techniques, as is the case of the authors Fontalvo et al. (2023), who used clustering analysis, artificial intelligence and data from the National Police between 2018 and 2022, in order to study crimes in Colombia. In that study, crime profiles were identified in different regions, allowing discerning areas with high levels of criminal acts. In addition, a double-layer neural network was developed with a high classification and prediction capacity of 97.7% for crime types and impacts. Likewise, Barragán et al. (2023) use artificial intelligence to predict crimes in space and time in Colombia. Demonstrating the relevance of the method for the evaluation of critical social events in Colombia.

Another study highlights the importance of predicting crime in developing countries by evaluating the effectiveness of predictive policing and different trainings (Galiani and Jaitman, 2023). Complementarily, authors Bi et al. (2023) highlight the need to predict legal sentences in criminal cases in a developing country and present the challenge of predicting prison sentences using a novel model that integrates judicial knowledge, allowing to significantly improve numerical legal predictions. From the above, the following questions arise: What is the appropriate method to characterize the different types of thefts evidenced in the 32 departments of Colombia, to identify critical areas? How to analyze the behavior of the different clusters of thefts presented in the 32 departments of Colombia in the year 2021, to establish crime trends? By means of what technique is it possible to forecast the theft clusters that will be in the 32 departments of Colombia in the future, for decision making?

The above approaches lead us to propose the following objectives for this research: i) Characterize the different types of thefts in the 32 departments of Colombia, to identify critical areas; ii) Analyze the behavior of the different theft clusters in the 32 departments of Colombia in 2021, to establish crime trends; and, iii) Forecast the theft clusters that will exist in the 32 departments of Colombia in the future, for decision making. Next, the theoretical foundation associated with the object of study of this research is presented, followed by the methodology developed to achieve the objectives. Subsequently, the levels of theft in the different departments of Colombia are characterized and finally the violence clusters are calculated, to finally establish a double-layer neural network to predict the future belonging of a city to a specific profile of violence established.

LITERATUR REVIEW

Data analytics and clustering

Data analytics represents a multidisciplinary and technological approach process dedicated to the exploration, cleansing and transformation of data sets for the purpose of identifying relevant patterns, trends and insights (García et al., 2021). This process encompasses multiple phases, from initial data acquisition to interpretation of results. During the acquisition stage, data are collected from various sources, such as databases, sensors, or logs, and undergo prior preparation for further analysis. Data cleaning encompasses the identification and correction of errors, outliers and missing data to ensure data integrity (Awudu et al., 2020). It is worth highlighting that, data analysis plays a critical role in various spheres, which include business, scientific, health and governmental fields (Mölder et al., 2021). Meanwhile, clustering is a technical process of grouping data based on the similarity between them. Its objective lies in organizing objects into groups (clusters) where the elements within the same group are more similar to each other than to those in other groups. This approach is widely used in data mining and analysis to discover hidden structures and patterns. While it is true, the clustering process involves selecting an appropriate algorithm that evaluates similarities between data using measures such as Euclidean distance or similarity coefficients.

Studies such as Rodríguez et al. (2020) address the interpretation of crime rates using data analytics, highlighting the time lag in reporting, which creates a benevolent bias when comparing partial periods. They analyzed data from 2005 to 2018 in Colombia, finding average lags in homicides of 2.85%, for theft at 11.8%, the same for personal injury of 12.7%, followed by 18.9% for domestic violence and finally, in sexual crimes of 30.5%. Likewise, Ordoñez et al. (2019) used data analysis to address the alarming growth of violent homicides in Latin America, especially in nations such as Mexico, Colombia and Venezuela. Data from the crime observatory of the Colombian National Police and the

Attorney General's Office were used to analyze the situation in Bogota in 2019, where 1,032 homicides were recorded at a rate of 14.3 per 100,000 inhabitants. Similarly, Maestre et al. (2023) research work uses data analysis with government information to examine shoplifting trends in Medellín (2014-2020). Seasonal patterns, preference for public places, and changes during the COVID-19 pandemic are highlighted, emphasizing the usefulness of open data in monitoring citizen action.

Artificial intelligence

Artificial intelligence (AI) has seen significant advances in recent decades, driven by the rapid development of technologies such as deep learning and big data processing (Khasawneh, 2023). In this context, AI's ability to analyze complex patterns and extract useful information has led to its application in a variety of areas, from automating industrial processes to improving healthcare and optimizing business decision-making (Shaqra, 2023). It is worth highlighting that, artificial intelligence refers to the ability of machines and software programs to learn, reason, adapt and make decisions in a manner similar to the human mind, using algorithms and mathematical models to process data and perform specific tasks (Davenport and Kalakota, 2019). As an illustration, in their article, Fontalvo et al. (2023) employ artificial intelligence and data analytics to address the classification and prediction of violent crime in Colombian departments between 2018 and 2022. Using clustering, they identified four crime clusters highlighting the most affected areas. Subsequently, they developed a neural network with a predictive capacity of 97.7%. This methodology is valuable for anticipating and addressing crime in Colombia, contributing significantly to effective crime prevention and response. Likewise, the article by Barragán et al. (2023) highlights the use of artificial intelligence to predict crimes in critical areas of Bogotá, given the growth of crime in the city. After reviewing 132 articles out of 3015 found in databases between 2019 and June 2021, it is concluded that neural network algorithms are highly effective in determining locations with crime levels. Similarly, artificial intelligence is highly beneficial in the field of data analytics by enhancing analysts' skills in terms of the speed and magnitude of data that can be processed (Miller et al., 2022). Since, artificial intelligence, through the use of algorithms, facilitates the ability to analyze significant volumes of information and extract meaningful information. In addition, artificial intelligence can provide timely answers to questions that previously could not be answered. By investing in artificial intelligence and data analytics, companies can make strategic decisions up to 5 times faster than their competitors.

Neural networks

Neural networks develop behavior similar to that of the human mind. Their main purpose is to recognize patterns and make decisions, as part of artificial intelligence. Gawlikowski et al. (2023) This is seen in the study by Bitzer et al. (2022), in which they use neural networks to analyze cases of sexual homicide, focusing on the collection and analysis of traces at the crime scene. They examined 230 sexual homicide cases from the SHielD database using neural network-based models. The results indicate that trace collection and analysis are influenced by crime scene-specific factors. Similarly, authors De la Hoz et al. (press) in their article developed and implemented an approach for the evaluation of quality criteria in services provided by health care providers (EPS) in Colombia. This methodology involved the combined use of multivariate analysis techniques and artificial intelligence, specifically, neural networks. The result obtained allowed the identification of distinct groups (clusters) representing different quality profiles in the services offered by the EPSs. Undoubtedly, neural networks play a crucial role in data analytics due to their ability to discover patterns and make predictions in large data sets. These networks are machine learning algorithms that model multiple data sets (Cardozo, 2023). They use units called artificial neurons that are connected to each other to transmit information Gawlikowski et al. (2023). Neural networks can process various types of information, such as images, text and voice, which makes them very versatile.

O-exclusive function (XOR gates)

Within the neural models we find the calculation of the o-exclusive function (XOR gates) of two variables (H1, H2). In which it is presented that, when H1 is different from H2 its classification pattern is "C2" or "1", while when H1 and H2 variables are the same, its identification is "C1" or "0" (Table 1) (Sri et al., 2021). This is due to the fact that it is impossible to find a line X1-X2 in the plane that allows separating the elements of both classes (Figure 1).

Table 1. Ranges of logic variables for o-exclusive problem (XOR logic gate).

X_1	X_2	Type of membership
0	0	0
0	1	1
1	0	1
1	1	0

Perceptron

In terms of functionality, in the perceptron for any E-type unit, it receives its input (n) in xk values (usually random), whereupon a weighted sum is calculated. In turn, at the binary output of the i-th unit E, with i=1, 2, 3 ..., m, the results can be between {0,1} or {-1,1} (Figure 2) (Repetur, 2019).

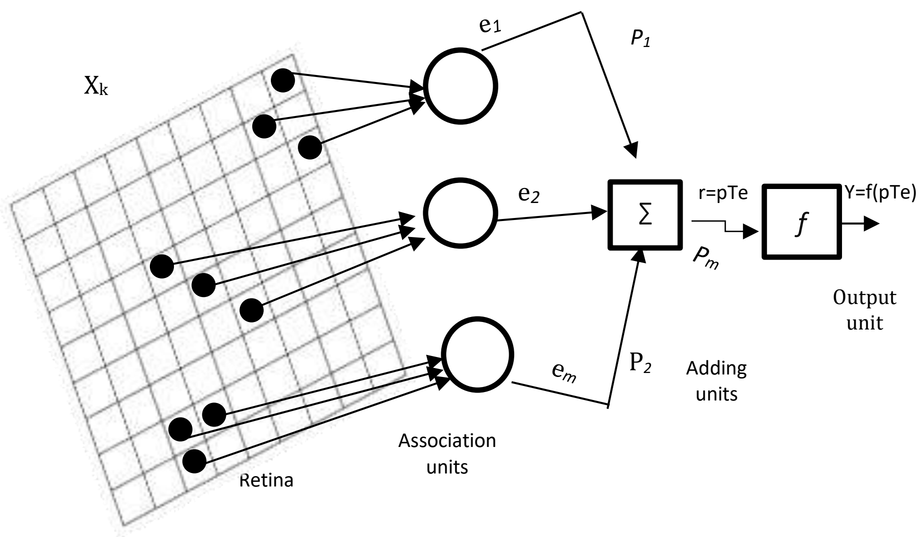

Figure 2. Original Perceptron.

Multi-input neural network

A multi-input neural network is a machine learning model that processes and relates information from multiple sources, allowing the capture of complex relationships and patterns in multidimensional data for prediction, classification or regression tasks (Delahoz et al., 2020). This approach is based on the integration of heterogeneous data, providing a robust and versatile framework for problem solving in various disciplines, where the synergy of multiple inputs allows addressing challenges more efficiently and accurately (Fontalvo et al., 2023). In turn, linkage measures are presented, which are based on the comparison of distances between pairs of elements to determine the proximity or relationship between them.

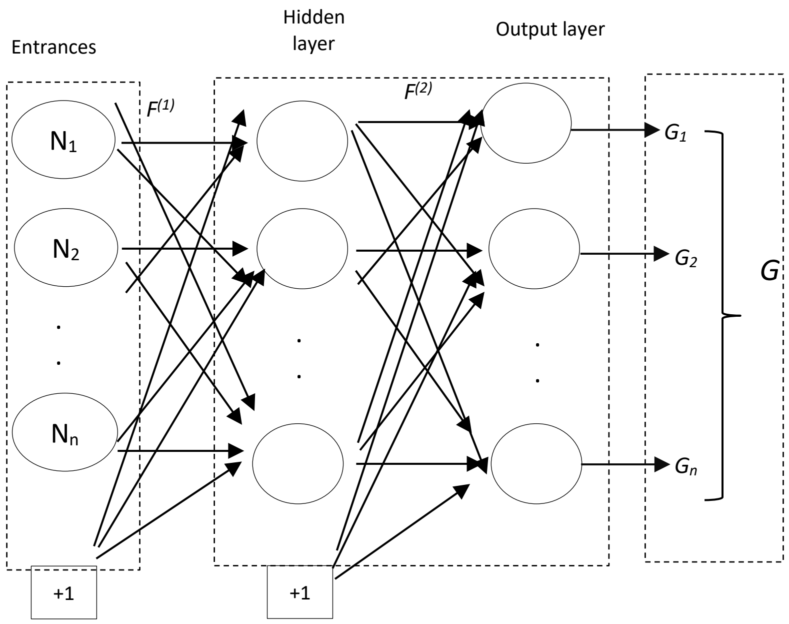

Figure 4. Structure of a multi-input neural network

Theft in Colombia

Colombia is a country that has historically been characterized by showing high crime rates, specifically different types of theft, which has persistently and negatively impacted social development, sustainable economy and the welfare and integrity of citizens in general (Coccia, 2018). This type of crime is defined by the criminal code of Colombia (2000) in its article 239 as the action of appropriating an object that belongs to another person with the intention of obtaining personal benefit or for the benefit of someone else. This criminal act can be classified into several categories, such as: i) simple theft, which involves the theft of a good without the use of force or violence; ii) aggravated theft, which occurs in the commission of the act with the use of threats or violence; iii) opportunity theft, which occurs when the occasion for the crime is developed without prior planning; and, iv) organized theft, which involves the participation of criminal groups to commit the act (Roca y Trespaderne, 2021). In recent years, several researchers such as Mandalapu et al. (2023) and Delahoz et al. (2020) have published articles using neural networks, artificial intelligence, clustering, among other methods to classify, assess and project violent crimes such as theft, discovering patterns and trends in the occurrence of these criminal acts in different countries and under different conditions. Likewise, in their work Esquivel, et al. (2020) presented a spatio-temporal prediction model of criminal events in Baltimore. Their approach integrated a convolutional network and a long- and short-term memory network to anticipate the presence of crimes such as robberies and thefts in upcoming days. This model accounted for spatial and temporal correlations in previous data, representing a fundamental advance in crime prediction. The effectiveness of the model was evaluated using standard metrics, such as accuracy and area under the receiver operating characteristic curve, supporting its practical applicability.

METHOD

The methodology employed consisted of a cross-sectional study based on a quantitative analysis, using historical information generated by the Colombian National Police on thefts that occurred in Colombia's 32 departments in 2021. In a first step, historical information on thefts in Colombia in the year 2021 was compiled, including details on their various modalities. Subsequently, these data were cleaned and unified to simplify their understanding and analysis. Next, 3 clusters of thefts in Colombia were elaborated, which served as a basis for the prediction of future incidents using a double-layer neural network. It should be noted that, for the establishment of the theft clusters, 16 cluster models were evaluated (Minitab software) combining 4 linkage methods (Ward, average, centroid and complete) and 4 metrics for distance measurement (Euclidean, Euclidean square, Manhattan and Pearson square), of which the Ward linkage and the Pearson distance measure were considered for this study. On the other hand, for the prediction process, 6 architectures or neural network models (SPSS software) were evaluated, from which a batch training with activation functions in the hidden layer of hyperbolic and sigmoid tangent was chosen, 37.5% (12) of the data were assigned for training, 34.4% (11) for testing and 28.1% (9) for the sample reserve. Six independent variables were also used: i) cattle theft, ii) theft from motor vehicles, iii) theft from establishments, iv) theft from motorcycles, v) theft from persons and vi) piracy. The theft clusters were also considered as a dependent variable. Figure 6 shows the methodology for the development of this research.

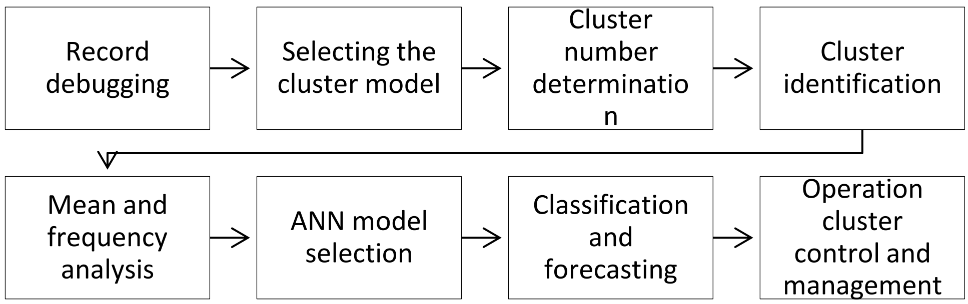

Figure 6. Schematic diagram of the study methodology

As previously mentioned, a data consolidation strategy was implemented to analyze the thefts in Colombia's 32 departments, covering the year 2021, with their different modalities (see Table 2).

Table 2. Characterization of the different types of theft in Colombia in 2021.

No	Department	Cattle theft	Theft of motor vehicles	Theft of establishments	Theft of motorcycles	Theft of persons	Piracy
1	Amazonas	0	0	70	32	72	0
2	Antioquia	124	1536	5782	7539	25388	20
3	Arauca	46	75	150	119	316	11
4	Atlántico	108	562	2057	2309	11177	30
5	Bolívar	166	120	1452	699	4941	4
6	Boyacá	132	34	819	172	2008	0
7	Caldas	29	16	593	104	1476	0
8	Caquetá	24	8	158	247	723	0

9	Casanare	169	23	455	280	1156	0
10	Cauca	40	454	902	2075	3167	11
11	Cesar	204	169	1435	819	2647	4
12	Chocó	2	2	164	337	390	0
13	Córdoba	107	20	505	476	1519	0
14	Cundinamarca	321	4530	13722	6194	92828	57
15	Guainía	1	0	49	3	52	0
16	Guaviare	11	0	41	37	54	0
17	Huila	119	47	800	1194	3925	0
18	Guajira	44	141	171	587	1464	0
19	Magdalena	115	49	594	428	2634	1
20	Meta	173	76	1186	791	3777	0
21	Nariño	20	208	607	855	4173	0
22	Norte de Santander	22	283	824	1018	3351	6
23	Putumayo	1	3	133	177	217	0
24	Quindío	33	22	597	203	1858	0
25	Risaralda	24	60	1315	282	2596	0
26	San Andrés	0	0	75	32	168	0
27	Santander	103	93	1727	1502	8247	5
28	Sucre	71	12	485	365	1543	1
29	Tolima	143	48	953	485	4396	5
30	Valle	80	2104	5261	4596	17214	17
31	Vaupés	0	0	23	1	12	0
32	Vichada	5	0	4	13	43	0
Total	-	2437	10695	43109	33971	203532	172

RESULTS

Establishment of theft clusters in Colombia

The data in Table 2 for the year 2021 reveal the typology of robberies in Colombia. Figures such as 2,437 cases of cattle theft, 10,695 cases of car theft, 43,109 cases of theft from establishments, 33,971 cases of motorcycle theft, 203,532 cases of theft from persons (the highest rate) and 172 cases of piracy stand out. These numbers are essential to understand the dynamics of crime in the country.

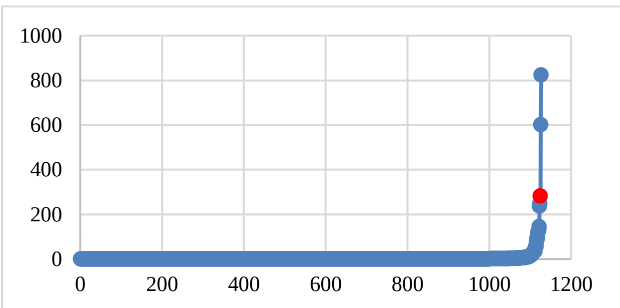

Figure 7. Elbow distances or diagram

On the other hand, Figure 7 shows the result of the clustering model with Ward linkage criterion and Manhattan distance measurement as the best for grouping the observations and in which it can be observed that the clusters are compact. Figure 8 presents a dendrogram that condenses the clustering of data related to theft in Colombia. The variables have been organized according to their similarity, starting with each variable as an individual cluster and gradually merging until they converge into a single grouping. This methodological approach has made it possible to identify theft patterns and analyze how they have evolved during the time period of interest.

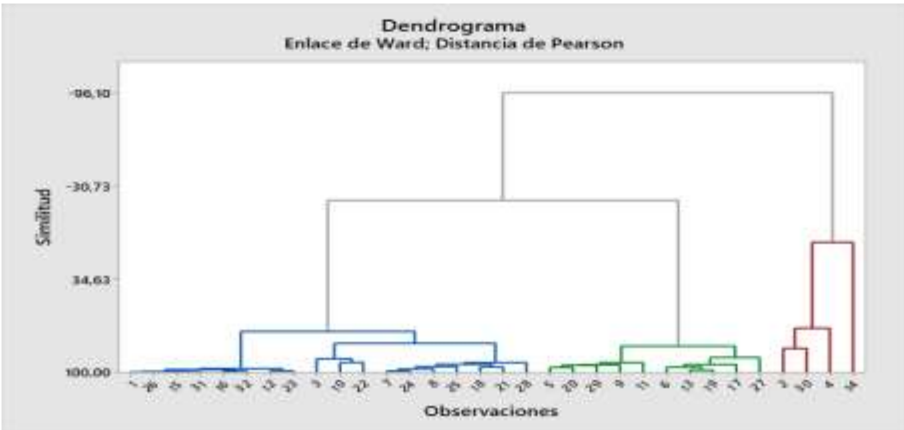

Figure 8. Dendogram of theft profiles in Colombia

As shown in the dendrogram in Figure 8, the analysis is based on three clusters of theft cases. The first cluster exhibits high homogeneity within its set of 18 observations, mainly in the departments of Amazonas and Santander corresponding to numbers 1 and 26, where there were only 3 types of theft out of the 6 under study and their crime levels are very low, ranging between 174 and 275, as well as among the departments of Choco, which only presented 2 cases of cattle theft and 2 cases of motor vehicle theft, and Putumayo with one case of cattle theft and 3 cases of motor vehicle theft, both identified as 12 and 23, respectively. The departments of Guainía with 105 thefts in the whole year 2021 and Vaupés with 36 cases of thefts, represented by numbers 15 and 31. Meanwhile, cluster 2, composed of the departments Cundinamarca

(117,652 thefts), Atlántico (16,243 thefts), Antioquía (40,389 thefts) and Valle (29,272), registers the highest levels of theft in Colombia, represented by the values 14, 4, 2 and 30, respectively. Likewise, these are the departments with the greatest differences in the number of thefts, as evidenced by the height of the branches in the dendogram. Finally, in the third dendogram the departments Bolivar (7,378 thefts), Meta (6,003 thefts), Tolima (6,030 thefts), Casanare (2,083 thefts), Cesar (5,278 thefts), Boyaca (3,165 thefts), Córdoba (2,627 thefts), Magdalena (3,821 thefts), Huila (6,085 thefts) and Santander (11,677 thefts), with differences between 1,400 and 6,000 cases of thefts and with high cases of thefts in all typologies.

Table 3. Theft clusters in the departments of Colombia

	Number of observations	Within the sum of squares of the cluster	Average distance from centroid	Maximum distance from centroid
Conglomerate 1	18	5,1368	0,45972	1,36730
Conglomerate 2	4	55,3633	3,35861	5,94386
Conglomerate 3	10	2,9917	0,51708	0,84756

When analyzing the final partitioning of the clusters in Table 3, it is highlighted that cluster 3 is the most homogeneous of all with a sum of squares at 2.9917, showing that the data points within the cluster are closer to each other. Furthermore, this cluster is considered to be more cohesive and better defined compared to the second one, since it has an average distance from the centroid of 0.51708, while the second cluster has a measure of 3.35861, showing that it is not compact. Likewise, cluster 2 is heterogeneous, since its data present a lot of difference, which is proven by a sum of squares of 55.3633. Finally, cluster 1 presents a high homogeneity, with a sum of squares of 5.1368, in spite of this, it does not exceed that of cluster 3. Even so, considering that it is the cluster that groups the most data, its average distance from the centroid of 0.45972, shows that it is the most compact.

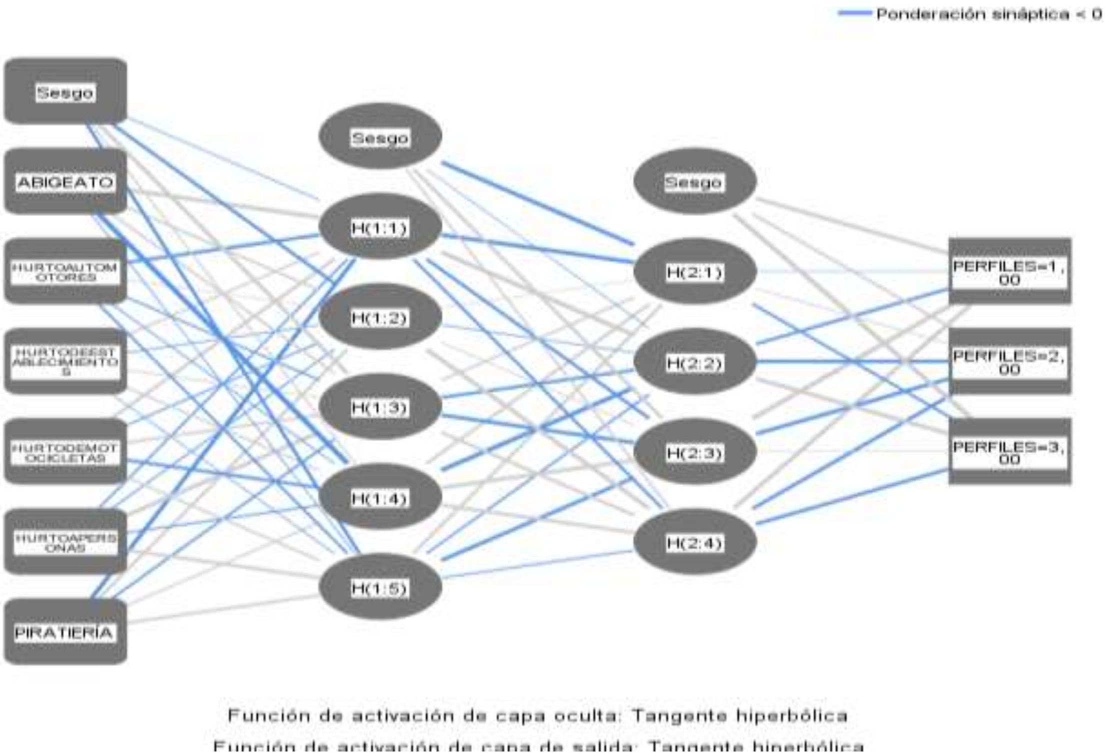

Figure 9. Estructura de la red neuronal de doble capa para el proceso de pronóstico de los perfiles de hurto.

Figure 9 presents the graphical representation of the dual-layer neural network designed to perform theft forecasting in the 32 departments of Colombia. This NR structure has been configured to achieve a considerably high level of predictive capacity.

Table 5. Prognostic capability of the proposed neural network model

Ranking					
		Forecast			
Sample	Observed	1,00	2,00	3,00	Percent correct
Training	1,00	7	0	0	100,0%
	2,00	0	1	0	100,0%
	3,00	0	0	4	100,0%
	Overall percentage	58,3%	8,3%	33,3%	100,0%
Tests	1,00	5	0	0	100,0%
	2,00	0	1	1	50,0%
	3,00	0	0	4	100,0%
	Overall percentage	45,5%	9,1%	45,5%	90,9%
Reservation	1,00	6	0	0	100,0%
	2,00	0	1	0	100,0%
	3,00	0	0	2	100,0%
	Overall percentage	66,7%	11,1%	22,2%	100,0%

Dependent variable: PROFILES

As can be seen in Table 5 related to the training, test and reserve samples in the prognosis, it can be observed that the accuracy was 100%, 90.9% and 100%, respectively. This results in an average capacity of 96.97% for the proposed neural network structure. This shows the relevance of the proposed neural network for predicting theft profiles in Colombia. As a significant finding of this research it can be stated that, of the three large conglomerates established, it can be evidenced that conglomerate 2, composed of the departments Cundinamarca (117,652 thefts), Atlántico (16,243 thefts), Antioquía (40,389 thefts) and Valle (29,272), registers the highest levels of theft in Colombia, represented with the values 14, 4, 2 and 30, respectively. In contrast to the above, cluster 1, which groups 18 departments, presents high homogeneity and is the most compact, with a sum of squares of 5.1368 and an average distance from the centroid of 0.45972. In addition, it should be noted that, within this conglomerate, the departments of Amazonas and Santander, corresponding to numbers 1 and 26, presented 3 types of theft out of the 6 under study and their crime levels are very low, ranging between 174 and 275, as well as the departments of Choco and Putumayo, which only presented 4 cases of cattle theft and car theft each. Similarly, the prediction accuracy of belonging to one of the three previously established clusters of the double-layer neural network used in this study is 96.97%, according to the number of thefts detected in the future. This mathematically proves the effectiveness of the methodology proposed to respond to the research objectives and to study social problems with great rigor in any national or international context.

DISCUSSION

In relation to this research, previous studies, such as that of Ramirez et al. (2019), applied neural networks for the classification of crime incidents related to public safety, using a dataset consisting of only five features, obtained from information provided by the National Police of the city of Chicago, with which they managed to achieve an accuracy rate of 87.84%. In this area of crime prediction, Ribeiro et al. (2022) carried out a prediction of homicides in urban centers by using machine learning techniques. They developed a model that incorporated generic crime incident data as well as temporal and spatial information. They then evaluated a total of 11 classification methods, which led them to find that the Random Forest-based model achieved an accuracy rate of 76% in predicting the occurrence or non-occurrence of homicides in the month following other recorded crimes, thus establishing a benchmark in solving this problem. Therefore, the results achieved in this research are more accurate, considering the fact that 6 variables were used and the level of accuracy of the proposed network was 96.97%, higher than the results of the two studies mentioned above.

In another context, Stec and Klabjan (2018) conducted a study that addressed the crime problem in the city of Chicago by using different neural network architectures, including Feed Forward Neural Networks, Recurrent Neural Networks, Convolutional Neural Networks and a combination of Recurrent Networks with Convolutions. Using data specific to the city of Chicago, the results indicated that the combined architecture of Recurrent Networks with Convolutions proved to be more effective in terms of prediction, achieving an accuracy of 75.6% for Chicago and 65.3% for Portland. From other perspectives, researchers Mohmed and Lujay (2023) have shown the importance of analyzing non-international and intra-country conflicts as in this research. And they raise the importance of establishing the categories, norms and stakeholders in each context to mitigate the levels of violence, which is similar to what is analyzed in this research where the regions and characteristics of a group of departments are categorized in terms of theft levels.

CONCLUSION

The most significant contribution of this research lies in the methodological proposal that allows: (i) Analyzing and cleaning large volumes of data with varying characteristics; (ii) Characterizing the different types of thefts that were evidenced in the 32 departments of Colombia in the year 2021; (iii) Analyzing the behavior of the theft clusters to establish crime trends; (iv) Establishing a double-layer neural network with an accuracy level of 96. 97% to forecast the theft clusters that will exist in Colombia in the future, which will be useful for decision making.

As a theoretical and methodological contribution, this research provides the scientific and business community with a framework that combines the conceptual approaches of clustering, artificial intelligence, data analysis, neural networks and theft in Colombia. It also established an effective methodology that integrates clustering with artificial intelligence to assess, classify and project theft clusters in Colombia's 32 departments. These results are essential to support decision making and the implementation of measures aimed at effectively combating theft crime in Colombia.

The practical focus and relevance of this contribution lies in its direct implication on current social security, in the search for an environment free of violence and vandalism that will reduce theft by citizens in the different departments that are the subject of this research.

As future research and given the versatility of this technique, the academic, scientific, business and governmental community is encouraged to adopt and adapt this methodology in various areas of prognosis for the analysis of social problems at regional, national and international levels, in order to classify and predict patterns of behaviour with a high degree of accuracy.

Bibliography

Abdullayev, I., Et al. (2023). Impact of international migration on internal state security, Migration Letters: an international journal of migration studies, 20(S4). doi: 10.47059/ml.v20iS4.3860.

Awudu, I., Et al. (2020). Application of big data copula-based clustering for hedging in renewable energy systems, International Journal of Revenue Management, 11 (4), pp. 237. doi: 10.1504/ijrm.2020.110633.

Barragán, H., Et al. (2023). Artificial intelligence and video surveillance in space-time crime prediction and detection: a systematic review, Crime Review, 65(1), pp. 11–25. doi: 10.47741/17943108.398.

Bi, S. et al. (2023). Judicial knowledge-enhanced magnitudeaware reasoning for numerical legal judgment prediction, Artificial intelligence and law, 31(4), pp. 773–806. doi: 10.1007/s10506-022-09337-4.

Bitzer, S., Et al. (2022). Sexual homicide and the forensic process: The decision-making process of collecting and analyzing traces and its implication for crime solving, Forensic science international, 340(111446), pp. 111446. doi: 10.1016/j.forsciint.2022.111446.

Cardozo, K. (2023). Aplicación de redes neuronales artificiales para el pronóstico de precios de café, Revista colombiana de tecnologías de avanzada, 1(39), pp. 113–117. doi:10.24054/rcta.v1i39.1403

Coccia, M. (2018). Economic inequality can generate unhappiness that leads to violent crime in society, International journal of happiness and development, 4(1), pp. 1. doi: 10.1504/ijhd.2018.090488.

Davenport, T., y Kalakota, R. (2019). The potential for artificial intelligence in healthcare, Future Healthcare Journal, 6(2), pp. 94–98. doi: 10.7861/futurehosp.6-2-94.

De la Hoz, E., Fontalvo, T., & Lopez, K. (press). Method to evaluate and project service quality profiles in health promotion entities in Colombia through multivariate analysis techniques and artificial intelligence, Int. J. Process Management and Benchmarking.

Delahoz, E., Fontalvo, T., y Mendoza, A. (2020). Definición de Perfiles Geográficos de hurto de automóviles. Caso Aplicado en Cartagena, Justicia, 25(37), pp. 99–108. doi: 10.17081/just.25.37.4215.

Efendi, A., Et al. (2023). Crowdfunding for the defense sector in Indonesia: challenges and opportunities, Migration Letters: an international journal of migration studies, 20(6). doi: 10.47059/ml.v20i6.3477.

Esquivel, N., Et al. (2020). Spatio-temporal prediction of Baltimore crime events using CLSTM neural networks, IEEE access: practical innovations, open solutions, 8, pp. 209101–209112. doi: 10.1109/access.2020.3036715.

Fiscalía General de la Nación. (2019). Rendición de cuentas 2018-2019, Bogotá D.C.: Fiscalía General de la Nación.

Flórez, J. (2021). The effect of mass incarceration on criminality in Colombia, International Journal of Criminal Justice and Social Democracy, 10(2). doi: 10.5204/ijcjsd.1644.

Fontalvo, T., Vega, M., y Mejía, F. (2023). Método de clustering e inteligencia artificial para clasificar y proyectar delitos violentos en Colombia, Revista científica General José María Córdova, 21(42), pp. 551–572. doi: 10.21830/19006586.1117.

Galiani, S., y Jaitman, L. (2023). Predictive policing in a developing country: Evidence from two randomized controlled trials, Journal of quantitative criminology, 39(4), pp. 805–831. doi: 10.1007/s10940-022-09551-y.

García, CW., Barón, E. y Sánchez, SK. (2021). La inteligencia de negocios y el análisis de datos en los procesos empresariales, Revista Científica de Sistemas e Informática, 1(2), pp. 38– 53. doi: 10.51252/rcsi.v1i2.167.

Gawlikowski, J. et al. (2023). A survey of uncertainty in deep neural networks, Artificial intelligence review, 56(S1), pp. 1513–1589. doi: 10.1007/s10462-023-10562-9.

Guevara, J., y Jiménez, E. (2020). Derecho a la libertad y aplicación de la prisión preventiva en delitos de robo y hurto, Universidad Técnica de Ambato, Unidad de Posgrado, Maestría en Derecho Constitucional.

Khasawneh, M. (2023). The potential of Ai in facilitating cross-cultural communication through translation. Journal of Namibian Studies, 37, 107–130. doi: 10.59670/jns.v37i.4654

Maestre, G., Et al. (2023). Análisis de datos sobre los hurtos en la ciudad de Medellín desde un enfoque descriptivo, Revista de Investigación Desarrollo e Innovación, 13(1), pp. doi: 10.19053/20278306.v13.n1.2023.16059.

Mandalapu, V., Et al. (2023). Crime prediction using machine learning and deep learning: A systematic review and future directions, IEEE access: practical innovations, open solutions, 11, pp. 60153–60170. doi: 10.1109/access.2023.3286344.

Miller, T., Et al. (2022). Special issue on Explainable Artificial Intelligence (XAI), Artificial intelligence, 307(103705), pp. 103705. doi: 10.1016/j.artint.2022.103705.

Mohmed, H., and Lujay T. (2023). Definition of Internationalized Armed Conflicts and their Legal Nature, Migration Letters 20(S5), pp. 636–645, doi: 10.47059/ml.v20iS5.4042

Mölder, F., Et al. (2021). Sustainable data analysis with Snakemake, F1000Research, 10, p. 33. doi: 10.12688/f1000research.29032.2.

Ordoñez, H., Pardo, C., y Cobos, C. (2019). Detección de tendencias de homicidios en Colombia usando Machine Learning, Revista facultad de ingeniería, 29(54), pp. e11740. doi: 10.19053/01211129.v29.n54.2020.11740.

Ramirez, U., Tello, E., and Mata, J. (2019). Predicting incidents of crime through LSTM neural networks in smart city domain, En Proc. SMART 8th Int. Conf. Smart Cities, Syst., Devices Technol, pp. 32–37.

Repetur, A. (2019). Redes Neuronales Artificiales, Universidad Nacional del Centro de la Provincia de Buenos Aires. Tesis pregrado, Repositorio institucional de la Universidad Nacional del Centro de la Provincia de

Buenos Aires.
Ribeiro, J., Et al. (2022). Prediction of homicides in urban centers: A machine learning approach, En Lecture Notes in Networks and Systems, pp. 344–61. Cham: Springer International Publishing.
Roca, C., y Trespaderne, A. (2021). El delito leve de hurto en la ciudad de Barcelona. Perfil del delincuente y comparación en función del tipo de hurto, Revista logos ciencia & tecnología, 13(2), pp. 39–55. doi: 10.22335/rlct.v13i2.1412.
Rodríguez, J., Et al. (2020). The analytical difficulty of the time lag in reporting and its relevance for the analysis of crime indicators in Colombia (2005-2018), Revista criminalidad, 62(3), pp. 49–63. doi: 10.47741/17943108.221.
Shaqra, R. (2023). The role of artificial intelligence in enhancing the leadership skills of academic leaders in public and private universities in the Northern Region from The Point of View of faculty members. Journal of Namibian Studies, 37, 290–320. doi: 10.59670/jns.v37i.4833
Sri R., Et al. (2021). Design and performance analysis of XOR and XNOR functions at low VDD using 130nm technology, journal of VLSI circuits and systems, 3(1). doi: 10.31838/jvcs/03.01.05.
Stec, A., and Klabjan, D. (2018). Forecasting crime with deep learning, En arXiv [stat.ML], 1. doi: 10.48550/ARXIV.1806.01486.