

Crime Prediction Using Machine Learning: A Systematic Review and Future Directions

Marcelo Adrian Arevalo-Cervera¹, Javier Gamboa-Cruzado², Carlos Diaz³,
Ángel Núñez Meza⁴, Alex Salazar-Marzal², Fabrizio Del Carpio Delgado⁵,
Mario Rodríguez Vasquez⁵, Cesar Jesús Núñez-Prado^{6,*}

¹ Universidad Nacional Federico Villarreal,
Peru

² Universidad Nacional Mayor de San Marcos,
Peru

³ Universidad César Vallejo,
Peru

⁴ Universidad Nacional Daniel Alcides Carrión,
Peru

⁵ Universidad Nacional de Moquegua, Facultad de Ingenierías,
Peru

⁶ Instituto Politécnico Nacional, ESIMEZ,
Mexico

2021018452@unfv.edu.pe, {jgamboac, alexmelecio.salazar}@unmsm.edu.pe, cdiazma@ucv.edu.pe,
anunezm@undac.edu.pe, {fdelcarpiod, mrodriguezv}@unam.edu.pe, cesar.jnprado@gmail.com

Abstract. The increase in urban crime has promoted the use of Machine Learning as a predictive approach to support crime prevention and decision-making in public safety. This systematic review aimed to analyze the application of Machine Learning in Crime Prediction by identifying technologies, application areas, editorial quality, conceptual foundations, and predominant thematic categories. The Kitchenham and Charters methodology was followed through searches in Scopus, IEEE Xplore, Web of Science, EBSCOhost, and ScienceDirect, with the search closed on July 12, 2025. After applying exclusion criteria and quality assessment, a final corpus of 60 papers was consolidated, and its data were organized and synthesized quantitatively and critically. The results show that Python was the predominant language, used in 75% of the studies, while applications were mainly concentrated in transportation, mental health, and public safety. Likewise, publications in Q1 and Q2 journals predominated, evidencing a corpus with adequate scientific visibility. Conceptually, Machine Learning was mainly defined from a technical perspective, focused on algorithms and computational

models. Finally, Cybercrime Forensics and Forensic Analytics emerged as motor themes, although several frequent lines still require greater theoretical and methodological consolidation.

Keywords. Deep learning, public safety, predictive model, crime prevention, bibliometric review.

1 Introduction

The increase in crime has motivated the use of Machine Learning (ML) to make crime predictions and improve public safety strategies, establishing a shift from reactive approaches to proactive data-driven models. The relevance of this area lies in its potential to reduce crime and optimize the allocation of police resources; however, knowledge remains fragmented: there is no consensus on the most effective algorithms,

results are often difficult to generalize, and multidisciplinary approaches that integrate social and technical perspectives are still lacking. Recent literature shows that Machine Learning-based Crime Prediction has focused, first, on improving the identification of critical areas, the allocation of police resources, and the anticipation of criminal patterns through preprocessing techniques, data mining, supervised learning, and feature selection, with high-accuracy results in contexts such as heat maps, criminal case analysis, and crime classification [1-3]. Likewise, the authors [4, 10, 11], in studies on urban crime and transfer learning, show that deep architectures, including deep neural networks, Bi-LSTM, and hybrid models, make it possible to capture complex patterns in large volumes of data, improving forecast quality compared with traditional approaches.

Complementarily, studies [5, 8, 9] compare and optimize predictive models such as LightGBM, Random Forest, linear regression, SVR, and hybrid algorithms, showing that model performance depends both on the nature of crime data and on the algorithm's capacity to handle nonlinear relationships, imbalanced data, and highly variable scenarios. Along the same lines, the authors [6, 7] demonstrate that model effectiveness depends not only on the selected algorithm, but also on its adaptation to the social, geographic, climatic, linguistic, or digital context, especially when limitations related to imbalanced data, low-resource languages, or particular territorial conditions exist.

Moreover, studies [12, 13, 18] emphasize that incorporating spatial, visual, contextual, and urban mobility data broadens the explanatory and predictive capacity of models by integrating information from geographic systems, urban images, police records, demographic data, and dynamic transportation flows. Continuing this perspective, the authors [14, 15] propose deep spatiotemporal architectures, such as 3D convolutional networks and CLSTM-NN models, which allow the temporal and territorial complexity of crime to be modeled, strengthening the prediction of criminal events in cities with a high density of historical information. Notwithstanding these advances, studies [16, 17, 19] warn that crime prediction faces relevant challenges in

intermediate cities, small areas, or data-scarce environments; therefore, graph-based approaches, heterogeneous information networks, and transferable models are strategic for improving territorial generalization. In turn, the authors [21-23], in studies focused on interpretable models, spatial analysis, and explainable artificial intelligence, show that the combination of XGBoost, SHAP, Max-p, and urban variables not only improves predictive accuracy, but also identifies explanatory factors associated with the built environment, land use, and the optimal location of surveillance devices. Finally, studies [20, 24, 25] extend the scope of crime prediction toward operational and digital scenarios by integrating criminology, computer vision, cybersecurity, public data, and balancing techniques such as SMOTE; together, these contributions show that the field is moving toward more robust, explainable, and adaptable models, although limitations remain regarding data availability, scalability, algorithmic fairness, and transfer to diverse geographic contexts.

Recent reviews show that Machine Learning approaches applied to crime and cybercrime prediction rely mainly on supervised models and traditional predictive techniques, such as random forest, support vector machine, logistic regression, and other classifiers, which have shown high levels of performance in cybercrime, video surveillance, and crime prediction contexts [62, 63, 66]. Likewise, the authors [74, 65, 69], in studies focused on video surveillance, artificial intelligence, and crime prediction, emphasize that Deep Learning architectures, together with technologies such as computer vision, the Internet of Things, edge computing, CNNs, LSTMs, GANs, and GNNs, are expanding automated analysis capacity, although limitations related to reliability, social acceptance, interpretability, and real-time data integration still persist.

Complementarily, studies [68, 71, 78] show that the field has evolved toward broader bibliometric and systematic approaches aimed at identifying trends, collaboration networks, thematic clusters, applied AI methods, and methodological challenges, including dependence on labeled data and the need to support police, scientific, and institutional decisions. In an applied line, studies [69, 75, 78] show that the use of Machine Learning

goes beyond urban crime prediction and extends to domains such as credit risk, crime prevention, and data-driven decision-making, confirming its potential to optimize processes, strengthen safety, and guide future research in real-world scenarios. Overall, the reviews analyzed confirm that the field is moving toward more sophisticated, hybrid, and explainable models; however, it still needs to overcome challenges related to data quality, the availability of labeled information, algorithmic transparency, empirical validation, and the transfer of models to different geographic, institutional, and operational contexts [65, 69, 78].

Although the literature shows relevant advances in the use of Machine Learning for Crime Prediction, knowledge remains fragmented in technical, methodological, and contextual terms. There is still insufficient consensus on which algorithms, languages, techniques, or predictive approaches provide better results according to the type of crime, data quality, territorial scale, or application context. In addition, many studies report high accuracy but present limitations in generalizing their findings to other urban, rural, or institutional environments.

Gaps also persist regarding interpretability, transparency, reliability, data availability, and social acceptance of models, especially when they may influence sensitive public safety decisions. This research is justified by the need to systematize and integrate scientific evidence on Machine Learning applied to Crime Prediction, considering its technical, bibliometric, conceptual, and thematic dimensions. Unlike reviews focused only on algorithms or specific applications, this study makes it possible to identify programming languages, application areas, quartile levels, theoretical foundations, and predominant thematic categories. Consequently, it provides a comprehensive view of the field, facilitates the identification of trends and gaps, and offers a rigorous basis for guiding future research and applications in public safety, urban management, cybersecurity, and data-driven decision-making.

The objective of this systematic review is to analyze the effectiveness and accuracy of Machine Learning applications aimed at Crime Prediction, based on previous scientific studies. The paper is organized as follows: Section 2 presents the theoretical Background; Section 3 describes the

review method, including research questions, sources, search strategies, exclusion criteria, PRISMA selection, quality assessment, data extraction, and synthesis of findings; Section 4 presents the results and discussion; and Section 5 develops the conclusions and future lines of research.

2 Background

This section presents the background of Machine Learning and Crime Prediction based on all the studies analyzed for this systematic literature review.

2.1 Machine Learning

Machine Learning comprises a set of artificial intelligence techniques that use algorithms to analyze data, identify complex patterns, including spatiotemporal trends, and build predictive models, such as neural networks and classification models, aimed at anticipating future events, including the occurrence of crimes or the evolution of crime rates [1, 16, 34, 37]. In this context, these techniques apply predictive algorithms, such as regression models, SLR, MLR, and DTR, classification methods, and supervised or unsupervised approaches, with the purpose of processing historical crime data, recognizing criminal patterns, estimating risk areas, and forecasting specific incidents [8, 27, 41]. Complementarily, studies applied to sentiment analysis argue that Machine Learning makes it possible to generate intelligent algorithms capable of learning from data without relying exclusively on rule-based programming, which reinforces its usefulness for classification and automated analysis tasks [61].

2.2 Crime Prediction

According to several authors [3, 9, 18, 25, 32], crime prediction is based on the analysis of historical data and contextual variables, including demographic, socioeconomic, environmental, and dynamic factors, through machine learning algorithms, with the purpose of identifying spatiotemporal patterns and associated risk

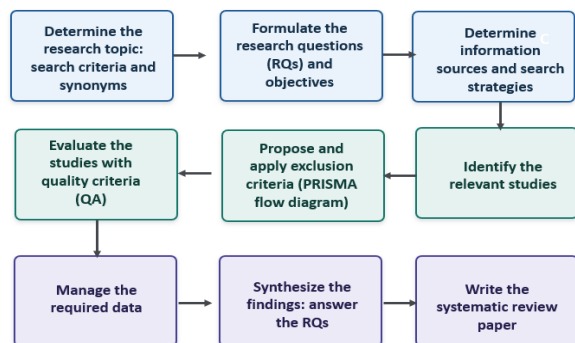


Fig. 1. Stages of the SLR

factors, such as poverty, education level, or urban mobility. In this field, models such as linear regression, decision trees, and random forests are used to estimate crime rates at different levels of territorial and temporal aggregation. Likewise, other authors [4, 11, 49] have proposed deep learning-based approaches, including deep neural networks, BiLSTM, and ConvBiLSTM, integrated with techniques such as SMOTE, hyperparameter tuning, and social media sentiment analysis, to improve predictive accuracy. Complementarily, Ribeiro and colleagues [38] apply graph convolutional neural networks, specifically GraphSAGE, to analyze criminal networks and predict criminal associations, illicit flows, and recidivism, achieving high levels of accuracy.

2.3 Machine Learning as an Analytical Foundation for Crime Prediction

Machine Learning constitutes a central methodological axis for crime prediction, since it makes it possible to transform large volumes of historical, spatial, temporal, and contextual data into computationally interpretable patterns capable of anticipating the occurrence, intensity, and territorial distribution of criminal events. This relationship is supported by supervised, unsupervised, and deep learning approaches, which, through classification algorithms, regression, neural networks, spatiotemporal models, and optimization techniques, make it possible to capture complex associations among social, environmental, and behavioral factors; however, its application faces challenges related to data quality and bias, model interpretability,

privacy, the availability of georeferenced information, and the risk of reproducing structural inequalities. Consequently, the integration between Machine Learning and Crime Prediction drives the evolution toward more robust, explainable, and adaptive predictive paradigms, making it possible to connect this conceptual basis with the methodological and bibliometric analysis developed in the following sections of the study.

3 Review Method

The purpose of a systematic literature review is to rigorously synthesize the available knowledge on a specific topic in order to obtain generalizable findings that support new lines of research, as stated by Kitchenham and Charters [77]. In this study, the methodology proposed by these authors was adopted, and adaptations aimed at strengthening the traceability and transparency of the process were incorporated, such as tables of search terms, matrices of search equations, a PRISMA diagram, and matrices for consolidating results. Similarly, previous systematic reviews have combined the Kitchenham and PRISMA guidelines to organize search sources, exclusion criteria, study selection, quality assessment, data extraction, and synthesis of results [79]. The development of the applied methodological procedure is presented in Figure 1.

3.1 Formulation of the Research Questions and Objectives

Conducting a systematic literature review requires formulating clear, precise research questions aligned with the purpose of the study, since these guide the search, selection, data extraction, and analysis of scientific evidence. In this study, the research questions were defined to structure the methodological process. These questions are presented in Table 1.

3.2 Information Sources and Search Strategies

To ensure the methodological rigor of the systematic literature review, academic sources of recognized scientific reliability were selected,

Table 1. Research questions and objectives

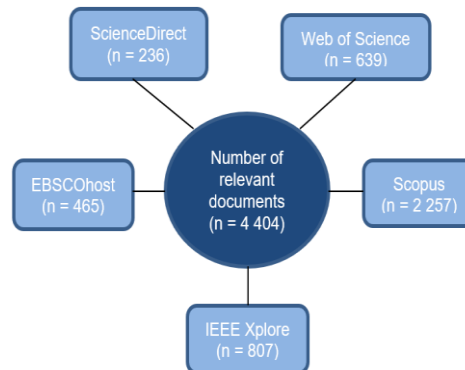
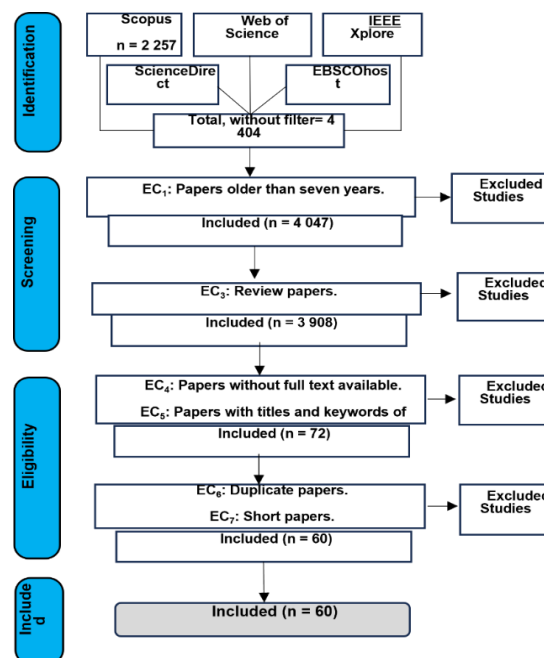
Question	Objective
RQ1: Which programming languages are most frequently used in the development of Machine Learning-based solutions?	Identify the programming languages most widely used in the development of Machine Learning models.
RQ2: In which business areas are Machine Learning models predominantly applied?	Identify the main business areas in which Machine Learning models are being applied.
RQ3: What quartile levels are presented by the scientific journals in which studies on Machine Learning and Crime Prediction have been published?	Examine the quartile levels of the scientific journals that have published research related to the application of Machine Learning for Crime Prediction.
RQ4: What conceptual definitions, theoretical foundations, and theoretical models have been reported in studies on Machine Learning?	Analyze the definitions, theoretical foundations, and theoretical models with the greatest presence in studies on Machine Learning.
RQ5: Which keywords present frequent co-occurrence patterns in research on Machine Learning and Crime Prediction?	Obtain the keywords that most frequently co-occur in research on the use of Machine Learning and Crime Prediction.
RQ6: Which thematic categories predominate in research on Machine Learning and Crime Prediction?	Classify research on Machine Learning according to the thematic categories most frequently addressed in Crime Prediction.

Table 2. Search terms and their synonyms by conceptual group

Conceptual Group	Descriptor
Machine Learning	machine learning/ml/predictive model/deep learning/ artificial intelligence
Crime Prediction	crime prediction/public safety/crime prevention/citizen safety

Table 3. Information Source and Search Equations

Source	Search Equation
Scopus	TITLE-ABS-KEY(("machine learning" OR "ml" OR "predictive model" OR "deep learning" OR "artificial intelligence") AND ("crime prediction" OR "public safety" OR "crime prevention" OR "citizen safety"))
IEEE Xplore	("Abstract": "machine learning" OR "Abstract": "ml" OR "Abstract": "predictive model" OR "deep learning" OR "Abstract": "artificial intelligence") AND ("Abstract": "crime prediction" OR "Abstract": "public safety" OR "Abstract": "crime prevention" OR "Abstract": "citizen safety")
Web of Science	(AB= ("machine learning" OR "ml" OR "predictive model" OR "deep learning" OR "artificial intelligence") AND ("crime prediction" OR "public safety" OR "crime prevention" OR "citizen safety"))
EBSCOhost	AB ("machine learning" OR "ml" OR "predictive model" OR "deep learning" OR "artificial intelligence") AND ("crime prediction" OR "public safety" OR "crime prevention" OR "citizen safety")
ScienceDirect	TITLE-ABS-KEY(("machine learning" OR "ml" OR "predictive model" OR "deep learning" OR "artificial intelligence") AND ("crime prediction" OR "public safety" OR "crime prevention" OR "citizen safety"))

**Fig. 2.** Number of Results by Source**Fig. 3.** Develops the following phases: Identification, Screening, Eligibility, and Inclusion

including Scopus, IEEE Xplore, Web of Science, EBSCOhost, and ScienceDirect.

The comprehensive search in all databases was completed on July 12, 2025, a date that defined the temporal limit of the study. Therefore, 2025 records should be analyzed with caution because they correspond to an incomplete annual period.

The terms used to retrieve the relevant studies, as well as the search equations adapted to each

database, are presented in Tables 2 and 3, respectively.

3.3 Identified Studies

The search conducted in the previously defined information sources identified a total of 4,404 studies potentially relevant to the systematic literature review. The distribution of the results obtained from each database is presented in Figure 2.

Table 4. Quality assessment score

Ref.	QA1	QA2	QA3	QA4	QA5	QA6	Score
[1]	3	3	3	2	3	3	17
[2]	2	3	1	2	3	3	14
[3]	3	2	2	1	2	2	12
[4]	2	3	1	3	2	2	13
[5]	2	1	3	1	2	2	11
[6]	2	2	1	3	2	3	13
[7]	3	3	1	2	2	2	13
[8]	1	2	2	3	2	2	12
[9]	3	3	3	2	3	3	17
[10]	3	2	3	3	3	3	17
[11]	3	3	3	2	3	3	17
[12]	3	3	3	2	3	3	17
[13]	3	3	3	2	3	3	17
[14]	2	1	3	3	3	2	14
[15]	2	2	2	3	3	3	15
[16]	2	3	1	3	2	2	13
[17]	2	3	3	3	3	2	16
[18]	3	3	1	2	2	2	13
[19]	1	2	2	3	2	2	12
[20]	3	3	3	2	3	3	17
[21]	3	3	3	2	3	3	17
[22]	3	3	3	2	3	3	17
[23]	3	3	3	2	3	3	17
[24]	3	3	3	2	3	3	17
[25]	2	2	1	3	2	3	13
[26]	2	3	1	3	3	3	15
[27]	1	2	2	3	2	2	12
[28]	2	1	3	1	2	2	11
[29]	3	3	3	2	3	3	17
[30]	3	1	1	3	3	3	14
[31]	2	1	3	3	3	2	14
[32]	2	1	3	3	3	2	14
[33]	1	2	2	3	2	2	12
[34]	2	2	2	3	3	3	15
[35]	2	1	3	3	3	2	14
[36]	3	3	3	2	3	3	17
[37]	3	3	3	2	3	3	17
[38]	3	3	3	2	3	3	17
[39]	2	1	3	1	2	2	11
[40]	2	2	1	3	2	3	13
[41]	2	1	3	3	3	2	14
[42]	1	2	2	3	2	2	12
[43]	2	1	3	1	2	2	11
[44]	2	2	1	3	2	3	13
[45]	3	3	3	2	3	3	17
[46]	3	3	3	2	3	3	17
[47]	3	3	3	3	3	2	17
[48]	3	3	3	2	3	3	17
[49]	2	2	1	3	3	3	14
[50]	2	3	1	3	2	2	13
[51]	2	3	3	3	3	2	16
[52]	2	2	3	2	3	3	15
[53]	2	2	3	2	3	3	15
[54]	2	2	3	1	3	2	13
[55]	2	2	1	3	2	2	12
[56]	3	3	3	2	3	3	17
[57]	2	3	3	3	3	3	17
[58]	3	2	3	3	3	3	17
[59]	2	2	3	3	3	1	14
[60]	2	3	2	2	3	3	15

3.4 Exclusion Criteria

The criteria established for this research are as follows:

- EC1: Papers older than seven years.
- EC2: Papers not written in English.
- EC3: Review papers.
- EC4: Papers without full text available.
- EC5: Papers with titles and keywords of limited relevance.
- EC6: Duplicate papers.
- EC7: Short papers.

The study selection process is an important phase in a systematic literature review, where the selection criteria established in the previous point are applied. Following the PRISMA protocol, Figure 3 develops the following phases: Identification, Screening, Eligibility, and Inclusion.

3.5 Quality Assessment

After the rigorous application of the selection criteria, 60 papers were obtained that met the methodological standards defined for this review. To systematically assess their quality, a protocol based on six assessment criteria was applied: clarity in the formulation of the research objectives (QA1), methodological description and justification of the techniques applied (QA2), relevance of the data collection methods (QA3), analysis of the reliability of the results obtained (QA4), correspondence between the answers and the research questions (QA5), and contribution of the study to the scientific and business community (QA6). This approach is consistent with recent systematic reviews, in which quality assessment is aimed at verifying the clarity of the objective, methodological adequacy, empirical support for findings, thematic relevance, and coherence among results, conclusions, and research objectives [76]. The scoring system considered three levels: optimal (3 points), acceptable (2 points), and deficient (1 point), so that the sum of the six criteria determined the total score of each paper, with 11 points established as the minimum inclusion threshold. Table 4 presents the results of the quality assessment; since the 60 papers

reached or exceeded the established minimum score, all met the defined criteria and were retained for the following phases of analysis.

As observed, all assessed papers reached or exceeded the minimum score of 11 points; therefore, they met the established quality criteria and were retained in the final corpus of the review.

3.6 Data Extraction Strategy

From the selected papers, the relevant information was extracted to answer the proposed research questions, considering the following fields: title, URL, source, year, country, ISSN, publication type, publication name, authors, affiliations, quartile, H-index, abstract, keywords, methodological details, conclusions, abstract synthesis, and number of citations. It should be noted that not all papers addressed each research question with the same level of depth, an aspect reflected in their respective quality scores. To ensure orderly and efficient information management, Mendeley Desktop was used as the bibliographic manager, which facilitated the organization, classification, and administration of the references included in the review. Likewise, RAj, Research Assistant j, an artificial intelligence tool developed by Dr. Javier Gamboa-Cruzado, was used to support the processing, systematization, and analysis of the collected data, strengthening the traceability, consistency, and analytical depth of the study.

3.7 Synthesis of Findings

The results obtained for each research question (RQ1-RQ6) were organized in tables and represented through quantitative data, which made it possible to develop a comparative analysis among the studies included in the review. This procedure facilitated the identification of recurring patterns, predominant trends, and relevant lines of research concerning the use of Machine Learning for crime prediction.

4 Results and Discussion

This section presents and analyzes the results obtained from the selected papers, in correspondence with the objectives of the study

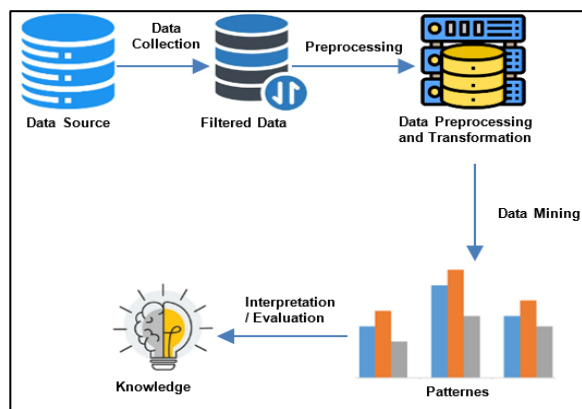


Fig. 4. Processing of selected documents

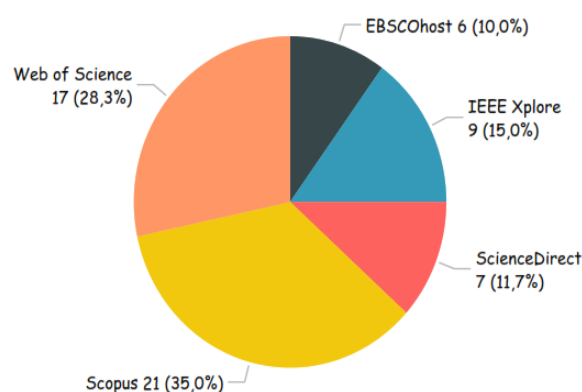


Fig. 5. Distribution of papers by source

and the related scientific literature. The documents were reviewed through a combined process of manual analysis and automated processing, which made it possible to organize, compare, and synthesize the evidence systematically.

The procedure applied for the automated processing of the selected documents is shown in Figure 4.

4.1 Overview of the Studies

Figure 5 presents the distribution of the selected papers according to the source database used in the systematic review. Its inclusion makes it possible to contextualize the documentary composition of the corpus and support the traceability of the academic sources consulted.

As shown in the findings, Scopus concentrates the largest proportion of papers, with 21 records (35.0%), followed by Web of Science, with 17 records (28.3%), evidencing the predominance of multidisciplinary databases with broad coverage in retrieving literature on Machine Learning and crime prediction. According to the observed distribution, IEEE Xplore contributes 9 papers (15.0%), confirming the presence of a relevant technical component associated with algorithms, computational models, and artificial intelligence approaches. Based on the identified pattern, ScienceDirect and EBSCOhost show lower participation, with 7 (11.7%) and 6 papers (10.0%), respectively, suggesting a complementary contribution from sources oriented toward applied, interdisciplinary, and social areas.

Cabanillas-Carbonell and colleagues [63] identified Scopus and ProQuest as the sources with the greatest representation, jointly concentrating 24% of the studies analyzed; in contrast, ScienceDirect registered the lowest contribution, with only 7% of the total.

Similarly, Esan and Chimbo [66] reported that Web of Science (WoS) was the most relevant data source, with 38 studies, while the Association for Computing Machinery (ACM) showed the lowest presence, with 13 studies. In the case of Iqbal and other authors [68], Google Scholar predominated as the main source, with 43 studies included; by contrast, IEEE showed minimal participation, with only one study.

Likewise, Jenga and collaborators [69] evidenced a different trend, since IEEE Xplore was the source with the greatest representation, with 29 papers, while ACM registered only 6.

Finally, Mandalapu and collaborators [72] indicated that most studies came from IEEE Xplore, although ScienceDirect had limited participation due to access restrictions; nevertheless, it should be noted that the authors did not specify the exact number of studies retrieved from each database, which limits the direct comparability of their results.

The concentration of papers in Scopus and Web of Science indicates that future reviews should prioritize multidisciplinary databases without excluding specialized sources, in order to broaden thematic coverage and reduce retrieval bias.

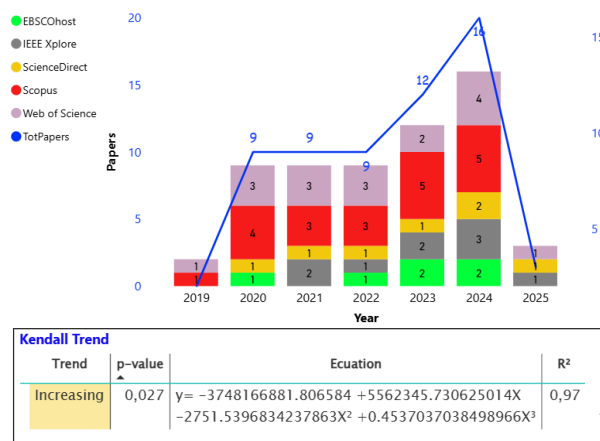


Fig. 6. Distribution of papers by year

This strategy can also be applied in sectors such as public health, transportation, banking, insurance, and risk management, where a balanced selection of sources is key to building robust and transferable evidence.

Figure 6 presents the temporal distribution of the selected papers during the period analyzed in the systematic review. Its inclusion makes it possible to contextualize the evolution of scientific production and support the trend analysis applied to the documentary corpus.

Based on the reported results, a progressive increase in papers is observed from 2019 to 2024, the year in which the highest value is reached, with 16 publications. Based on the evidence presented, the Kendall trend test confirms a statistically significant increasing behavior ($p = 0.027$), supported by a high fit of the temporal model ($R^2 = 0.97$), which reinforces the statistical consistency of the observed growth. According to the identified pattern, the reduction recorded in 2025 should be interpreted with caution because the comprehensive database search was completed on July 12, 2025 and, therefore, that year corresponds to an incomplete annual period.

Cabanillas-Carbonell and co-authors [63] reported that academic production reached its peak in 2023, with 14 published papers, representing a 250% increase compared with 2021, the year of lowest productivity, with only 4 publications. Consistently, Ersoz and colleagues [65] showed a marked variation in scientific

production, from the absence of publications in 2016 to a maximum of 30 papers in 2023, followed by 16 studies in 2024; consequently, the results suggest accelerated growth from 2020 onward, although with a subsequent reduction of 47% after the peak recorded in 2023. For their part, Cárdenas-Quispe and fellow researchers [67] emphasized that neural networks predominate in most studies, supported by deep learning theories that make it possible to identify nonlinear and complex relationships in large volumes of data; likewise, discriminant analysis, based on multivariate statistics, has a relevant presence, while comparative analysis fulfills a complementary methodological function by contrasting performance among different approaches.

In the study by Iqbal and collaborators [68], scientific production showed an upward trend between 2019 and 2022, with a progressive increase from 12 papers in 2020 to 19 publications in 2022; however, this growth was interrupted in 2024, when production decreased to 13 papers. In turn, Kounadi and co-workers [71] identified, for the 2001-2018 period, an irregular evolution of scientific production, with 2018 as the year of greatest activity, with 13 papers, while the rest of the period showed considerably lower figures, including years without publications; it should be noted that the only clearly recognizable trend was located between 2016 and 2018, when production rose from 1 study to 13 papers. Finally, Mandalapu and colleagues [72] reported that 2019 registered the lowest production, with 22 papers, while 2021 reached the maximum, with 51 publications; in contrast, 2022 showed a decrease to 29 papers, evidencing a trajectory of sustained growth until 2021, followed by a subsequent contraction.

The increasing trend evidences the recent consolidation of academic interest in the use of Machine Learning for crime prediction, which justifies expanding the analysis in later periods and in other geographic contexts. Likewise, incorporating statistical trend tests strengthens the methodological rigor of systematic reviews applicable to other data-intensive fields and predictive decision-making.

Table 5 presents the distribution of the 60 selected papers according to the publication journal, incorporating the number of papers,

Table 5. Distribution of papers by journal

N°	Publication Name	N° Papers	Citations	H Index	Quartile
1	IEEE Access	9	191	2610	Q1
2	ISPRS International Journal of Geo-Information	6	89	504	Q1
3	Computers, Materials and Continua	2	14	134	Q2
4	Electronics	2	20	220	Q2
5	Multimedia Tools and Applications	2	22	232	Q1
6	PLOS ONE	2	25	934	Q1
7	Sustainability	2	16	414	Q1
8	ACM Transactions on Spatial Algorithms and Systems	1	19	20	Q3
9	Applied Computational Intelligence and Soft Computing	1	14	26	Q1
10	Applied Network Science	1	10	32	Q1
11	Applied Spatial Analysis and Policy	1	4	33	Q2
12	Computación y Sistemas	1	0	163	Q3
...	...	...	...	...	...
	Total	60	731	6626	

Table 6. Distribution of publications and bibliometric indicators by country

N°	Country	N° Papers	% N° Papers	No. Citations	% No. Citations	H-Index	Citations/ Paper
1	US	11	11,1	167	12,3	893	15,2
2	China	10	10,1	61	4,5	879	6,1
3	UK	7	7,1	123	9,1	991	17,6
4	India	6	6,1	57	4,2	251	9,5
5	Pakistan	6	6,1	139	10,2	1520	23,2
6	Saudi Arabia	6	6,1	42	3,1	708	7,0
7	Malaysia	5	5,1	34	2,5	1251	6,8
8	Australia	3	3,0	14	1,0	468	4,7
9	Canada	3	3,0	39	2,9	457	13,0
10	Iraq	3	3,0	22	1,6	114	7,3
11	Korea	3	3,0	62	4,6	415	20,7
12	Bangladesh	2	2,0	39	2,9	310	19,5
13	Belgium	2	2,0	4	0,3	47	2,0
14	Jordan	2	2,0	6	0,4	134	3,0
15	Philippines	2	2,0	2	0,1	307	1,0
16	Poland	2	2,0	32	2,4	374	16,0
...	...	...	...	...	...	...	...
	Total	99	100,0%	1357	100,0%	12964	13,7

accumulated citations, H-index, and quartile of each source. Its inclusion makes it possible to assess not only the editorial concentration of the corpus, but also the bibliometric relevance and academic positioning of the journals supporting the review.

As shown in the findings, IEEE Access concentrates the largest number of papers, with 9

publications, 191 citations, an H-index of 2610, and Q1 classification, evidencing its dominant role as a dissemination channel for studies linked to artificial intelligence, machine learning, and computational applications.

According to the observed distribution, ISPRS International Journal of Geo-Information occupies a relevant position, with 6 papers, 89 citations, an

Table 8. Business Areas by Research Study

Business Area	Reference	Qty. (%)
Public safety	[1] [2] [4] [15] [20] [23] [26] [29-30] [32] [42] [44] [46] [48] [53-54] [56]	17 (25)
Smart Cities	[31] [52] [57]	3 (4)
Banking	[24]	1 (1)
Private Security	[23] [32]	2 (3)
Transport	[1] [8-10] [12] [15-17] [20] [23-24] [27] [36] [40] [42] [44] [48] [53] [59]	19 (28)
Technology and Cybersecurity	[7] [24] [32] [42-43] [52]	6 (9)
Mental health	[1-3] [7-8] [11] [15] [20] [24-25] [27] [36] [39] [42] [44-45] [49] [52] [54]	19 (28)
Real estate	[52]	1 (1)

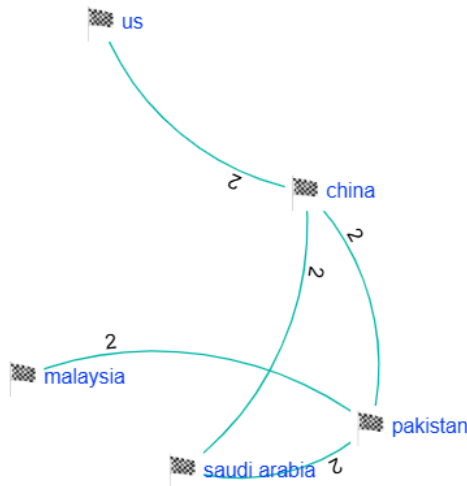


Fig. 7. Collaborative network among countries

H-index of 504, and Q1 quartile, reinforcing the importance of the geospatial component in crime prediction through models based on territorial data and urban patterns. Based on the identified pattern, the presence of Q1 journals such as

Multimedia Tools and Applications, PLOS ONE, and Sustainability, together with Q2 and Q3 journals, shows that scientific production is distributed among high-impact sources and specialized channels, reflecting the interdisciplinary nature of the field.

Kaur and Saini [78], when analyzing scientific production for the 2012-2022 period, identified Lecture Notes in Computer Science, including its subseries linked to artificial intelligence and bioinformatics, as the most productive source, with 104 published papers. Advances in Intelligent Systems and Computing ranked second, with 72 studies, while ACM International Conference Proceeding Series and IEEE Access shared third position, with 64 publications each. Overall, these results evidence a relevant concentration of scientific production in editorial series and sources associated with computer science, artificial intelligence, and intelligent systems.

This distribution suggests that future research on Machine Learning and crime prediction should consider journals with high bibliometric visibility, especially those located in Q1, without excluding specialized sources that may provide complementary methodological or thematic approaches. Likewise, the joint analysis of number of papers, citations, H-index, and quartile can be applied in other data-intensive sectors, such as public health, transportation, banking, insurance, urban management, and risk prevention, to guide publication strategies, scientific monitoring, and selection of relevant literature.

Figure 7 presents the collaborative network among the countries identified in the selected papers, while Table 6 organizes the distribution of publications and their main bibliometric indicators by country. Its inclusion makes it possible to jointly analyze the structure of international cooperation and the scientific weight of each country within the reviewed corpus.

As shown in the findings, scientific production is mainly concentrated in countries such as the United States, China, the United Kingdom, India, Pakistan, and Saudi Arabia, although bibliometric impact does not depend solely on the number of papers, since Pakistan registers 6 publications and reaches 139 citations, with 23.2 citations per paper, while the United States presents 11 papers and 167 citations, with 15.2 citations per paper; in

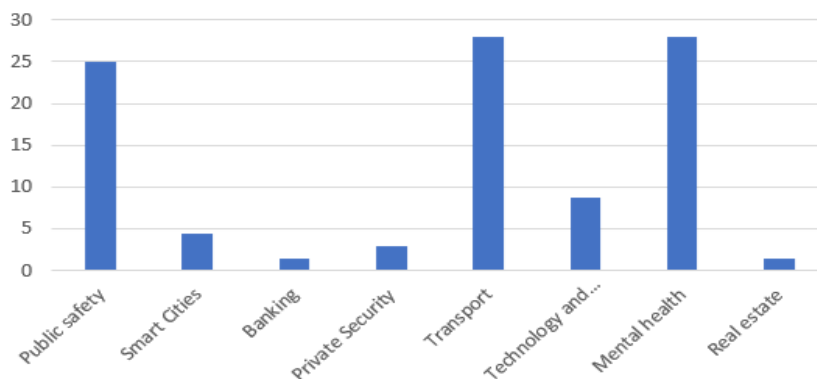


Fig. 9. Distribution of business areas



Fig. 10. Distribution of papers by quartile and country

addition, the collaborative network shows specific links among China, the United States, Saudi Arabia, Pakistan, and Malaysia.

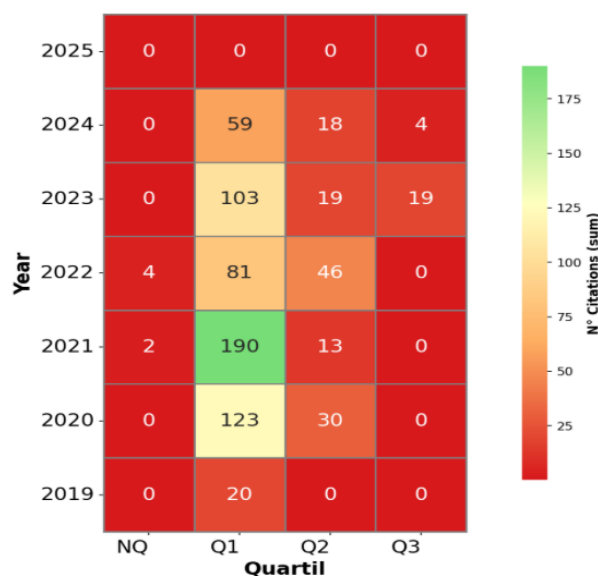
It suggests that research on Machine Learning and crime prediction combines relevant productive nuclei with patterns of international cooperation that are still selective and not highly dense.

4.2 Answers to the Research Problems

This section presents the answers to the formulated research questions, integrating the main results, their critical discussion, and possible projections for future studies. The findings are derived from a systematic bibliographic search on

Table 9. Distribution of papers by quartile and year

Quartile	2019	2020	2021	2022	2023	2024	2025	Total
Q1	2	6	6	5	6	9	3	37
Q2	0	3	2	2	4	5	0	16
NQ	0	0	1	2	1	0	0	4
Q3	0	0	0	0	1	2	0	3
Total	2	9	9	9	12	16	3	60

**Fig. 11.** Citations by year and quartile

Machine Learning and crime prediction, carried out in widely recognized academic sources and complemented by a rigorous process of selection, review, and quality assessment. As a result of this procedure, a final corpus of 60 papers was consolidated, considered relevant and methodologically adequate to support the analysis developed in this review.

4.2.1 RQ1: Which Programming Languages are Most Frequently Used in the Development of Machine Learning-Based Solutions?

Table 7 and Figure 8 present the distribution of the programming languages identified in the papers

that develop Machine Learning-based solutions, indicating the associated references and their relative frequency within the analyzed corpus. Its inclusion is relevant because it makes it possible to recognize the predominant technologies used to implement computational models and to assess the technical orientation of the reviewed studies.

As shown in the findings, Python concentrates a clearly dominant presence, with 21 records equivalent to 75%, evidencing its position as the main language for developing Machine Learning-based solutions. According to the observed distribution, Java reaches 4 records (14%), suggesting a secondary use, possibly associated with more structured development environments or applications that require integration with existing systems.

Based on the identified pattern, C++, C#, and PHP show marginal participation, with 1 record each (4%), indicating that their use in the analyzed studies was occasional and did not constitute a predominant trend within the reviewed corpus.

Dakalbab and colleagues [74] identified Weka as the most widely used tool for developing predictive models, with 11 studies (23%), followed by Python, with 8 studies (17%); in contrast, MATLAB, TensorFlow, and PyTorch registered marginal presence, with only one study each (2%). Similarly, Esan and Chimbo [66] reported that Weka also had the highest frequency of use, with 27%, while Python and R Studio reached an equivalent participation of 20% each; nevertheless, tools such as TensorFlow, MATLAB, and MALLET showed a minimal contribution, with 2% in all cases.

Together, both studies evidence that Weka maintains a relevant presence in the development of Machine Learning-based solutions, although Python appears as a recurrent and technically competitive alternative; therefore, the selection of tools seems to depend both on the type of predictive model and on the degree of accessibility, specialization, and ease of implementation required in each study. Studies on conversational technologies in e-Commerce also identify Python and Java as predominant languages for developing intelligent solutions, with 32% and 25%, respectively, which partially coincides with the centrality of Python observed in Machine Learning-based applications [70].

These results suggest that future research and business applications based on Machine Learning should prioritize technological ecosystems compatible with Python, without ruling out complementary languages when the implementation context requires integration with corporate platforms, legacy systems, or specific architectures. Likewise, this pattern can be contrasted in other sectors, geographic areas, and periods of time to assess whether Python's hegemony is maintained or whether alternative languages emerge according to new computational, regulatory, or business needs.

4.2.2 RQ2: In which Business Areas are Machine Learning Models Predominantly Applied?

Table 8 and Figure 9 present the distribution of the business areas in which the Machine Learning models identified in the reviewed papers are applied. Its inclusion is relevant because it makes it possible to recognize the sectors with the greatest application of the predictive approach and assess the practical orientation of the analyzed studies.

Based on the reported results, transportation and mental health concentrate the largest proportion of applications, with 19 records each (28%), evidencing that Machine Learning models are mainly used in domains where the anticipation of events, risks, or behaviors is critical for decision-making. Based on the evidence presented, public safety also shows a significant presence, with 17 records (25%), confirming its direct link with Crime Prediction, especially in contexts where it is necessary to analyze criminal patterns, mobility, incidents, and risk factors. According to the identified pattern, areas such as smart cities, technology and cybersecurity, private security, banking, and real estate show lower frequency, suggesting more specific or emerging applications within the reviewed corpus.

Kaur and Saini [78] showed that the Social Sciences concentrated the greatest academic production, with 3,090 publications, followed very closely by Computer Science, with 2,980, and Medicine, with 2,970; therefore, the results show that the application of Machine Learning models has a markedly interdisciplinary orientation by combining social, computational, and health-

related domains. In contrast, areas such as Business Management and Accounting, with 333 publications, and Environmental Science, with 234, registered a considerably lower contribution; nevertheless, it should be emphasized that their presence confirms the progressive expansion of these models toward business and applied fields where predictive analytics can support decision-making. Evidence on API and software deployment confirms that Machine Learning is also applied in business technology areas, especially for early error detection, monitoring, control, and improvement of organizational processes linked to information systems [73]. Together, these findings suggest that the use of Machine Learning is concentrated in areas with high data availability, complex problems, and a need for prediction, although its adoption in management and sustainability sectors still shows room for growth.

These results indicate that Machine Learning models have high transfer potential toward data-intensive business sectors, especially transportation, public safety, mental health, smart cities, banking, cybersecurity, and real estate management, where prediction can support operational, preventive, and strategic decisions. Likewise, the observed distribution justifies replicating this analysis in other geographic contexts and periods of time to determine whether the identified sectoral concentration responds to a global trend or to specific conditions of the studies included in this review.

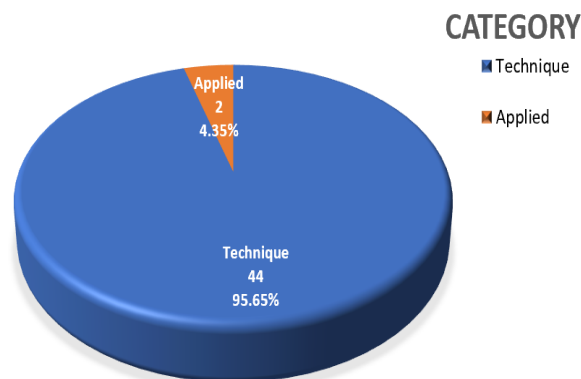
4.2.3 RQ3: What Quartile Levels are Presented by the Scientific Journals in which Studies on Machine Learning and Crime Prediction have been Published?

Figure 10, Table 9, and Figure 11 present the distribution of papers according to the quartile of the scientific journals, considering their relationship with country, year of publication, and number of accumulated citations. Their inclusion makes it possible to analyze the editorial quality of the corpus from a geographic, temporal, and bibliometric perspective, strengthening the interpretation of studies on Machine Learning and Crime Prediction.

Iqbal and colleagues [68] reported that 43% of the analyzed papers, equivalent to 24 studies, were published in Quartile 1 (Q1) journals, while

Table 10. Definitions by research area

Synthesized Definition	Category	Reference	Qty. (%)
ML uses algorithms, statistical models, and computational techniques to learn from data, identify patterns, and perform tasks automatically, without explicit programming.	Technique	[1-4] [6] [8-11] [13-15] [17] [19-22] [26-27] [29-30] [32] [34-35] [37-38] [41-51] [53] [55-60]	44 (96)
ML is applied to predict phenomena in specific domains, such as crime or geography, through models that generalize complex data-based relationships.	Applied	[12] [33]	2 (4)

**Fig. 12.** Distribution of definitions

27%, corresponding to 15 studies, were located in Q2 journals; therefore, most of the examined production is concentrated in editorial channels with higher academic positioning. In contrast, Q3 and Q4 journals showed a lower and equivalent participation, with 11% each, represented by 6 studies in both cases. Overall, these results suggest that research on Machine Learning and Crime Prediction tends to be published predominantly in higher-quartile journals, although

a smaller proportion of studies disseminated in medium- or low-impact sources persists, which should be considered when assessing the editorial quality and scientific visibility of the corpus.

This pattern evidences the need to prioritize high-impact sources in future reviews, while also considering countries and journals with lower presence when they provide contextual or methodological evidence relevant to understanding the application of predictive models in different institutional and geographic environments.

According to the observed distribution, papers published in Q1 journals predominate during the analyzed period, with 37 of the 60 records, followed by Q2 with 16, while NQ and Q3 show lower participation, with 4 and 3 papers, respectively; in addition, 2024 concentrates the highest number of publications, with 16 records, of which 9 correspond to Q1.

In the study by Iqbal and colleagues [68], 2022 registered the highest scientific production, with a total of 19 publications, concentrated mainly in Q1 journals (6 papers) and Q2 journals (5 papers), followed by Q3 (4 papers) and Q4 (1 paper). In contrast, the years with the lowest number of publications were 2020 and 2023, with 12 papers each; in both cases, most works were located in Q1 journals, while Q3 had the lowest relevance, with only one paper in each year.

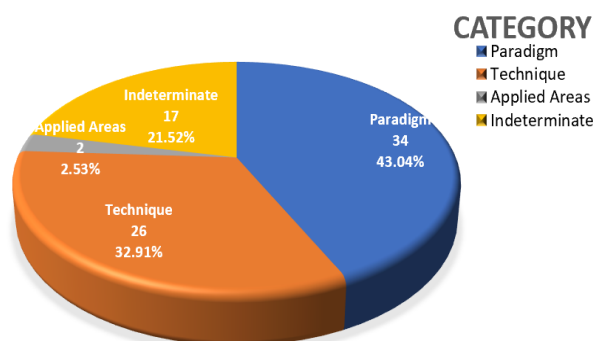
This composition supports the editorial strength of the corpus and makes it possible to use quartile as a complementary criterion to assess evidence quality in future systematic reviews, as well as to compare the evolution of the field in other periods, regions, and data-intensive domains.

According to the identified pattern, citations are mainly concentrated in Q1 journals, which accumulate the highest values in 2021 (190), 2020 (123), 2023 (103), 2022 (81), and 2024 (59), while Q2 presents a secondary and more moderate contribution, notably 46 citations in 2022 and 30 in 2020; in contrast, NQ and Q3 register low or null values in most years, evidencing that the bibliometric impact of the corpus is strongly associated with publications located in higher quartiles.

Iqbal and colleagues [68] stated that, in the United States, most studies were published in high-impact journals, with 17 papers in Q1 and 8 in

Table 11. Theoretical Foundations by research area

Synthesized Theoretical Foundation	Category	Reference	Qty. (%)
Conceptual approaches that support predictive and automated analysis models through artificial intelligence, machine learning, and computational learning.	Paradigm	[1-2] [6] [8] [11-12] [14-15] [20-21] [26-27] [29-30] [32-35] [37-38] [41-45] [47] [50-51] [53] [55-58] [60]	34 (43)
Quantitative and computational methods used to analyze data, model relationships, and make predictions in applied contexts.	Technique	[9] [12] [15] [19] [21] [26-27] [29-30] [32-33] [35] [37-38] [41-45] [47] [51] [53] [55] [57-58] [60]	26 (33)
Specific fields where intelligent models are implemented, such as predictive criminology and computer vision.	Applied Areas	[19-20]	2 (3)
Terms without clear classification into paradigms, techniques, or applied areas because of their ambiguity, generality, or lack of specificity.	Indeterminate	[3-4] [10] [13] [17-18] [22-23] [25] [28] [40] [46] [48-49] [52] [54] [59]	17 (22)

**Fig. 13.** Distribution of theoretical foundations

Q2, although publications in Q3 and Q4 were also recorded, with 1 and 4 papers, respectively. In contrast, India showed a lower presence, but

distributed between Q1 and Q4 journals, with 2 papers in each quartile. Likewise, countries such as Bangladesh, Argentina, Brazil, China, South Korea, and South Africa had more limited representation, with publications mainly concentrated in Q1, Q2, and Q3 journals; overall, these results evidence a geographic concentration of editorial impact in the United States, without disregarding the participation of other countries in channels of different scientific visibility.

This result suggests that, in addition to the number of published papers, the analysis of citations by quartile and year makes it possible to identify where the real scientific influence of the field is concentrated, so it can be applied in other sectors, such as public safety, transportation, health, banking, cybersecurity, and urban management, to guide publication decisions, bibliometric monitoring, and strategic literature selection.

4.2.4 RQ4: What Conceptual Definitions, Theoretical Foundations, and Theoretical Models have been Reported in Studies on Machine Learning?

Table 10 and Figure 12 present the classification of the conceptual definitions of Machine Learning reported in the reviewed studies, while Table 11 and Figure 13 organize the theoretical foundations identified according to analytical categories. Their inclusion makes it possible to examine how the literature conceptualizes Machine Learning and which theoretical foundations support its application in the context of Crime Prediction.

As shown in the findings, the technical definition of Machine Learning widely predominates, with 44 records (96%), evidencing that most studies conceive it as a set of algorithms, statistical models, and computational techniques oriented toward automatic learning from data, while the applied definition appears in only 2 records (4%), associated with its use in specific domains such as crime or geography.

The literature on e-Commerce and Android malware detection coincides in conceiving Machine Learning from a technical perspective, associated with algorithms and mathematical models capable of learning patterns from data, supporting predictive tasks, and executing classification, detection, and pattern recognition

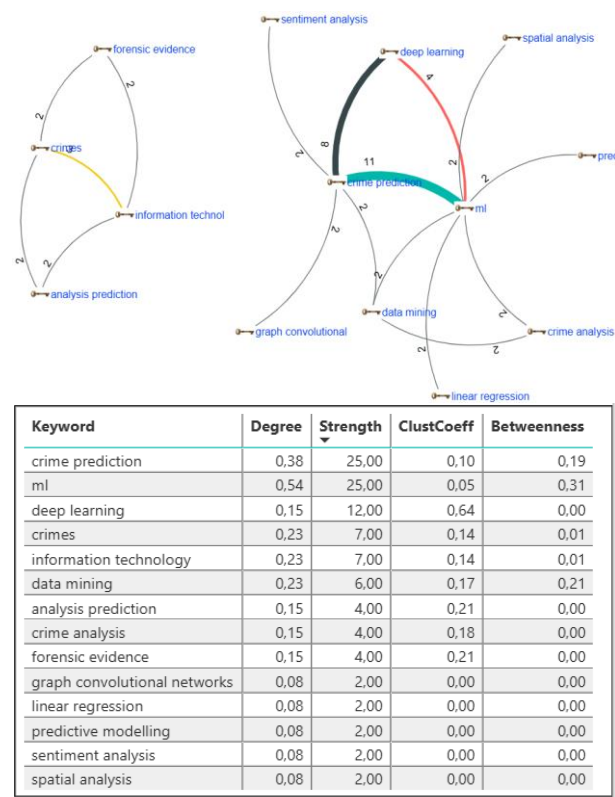


Fig. 14. Keyword co-occurrence

processes in complex and digital environments [64, 67].

This technical predominance suggests that future research should complement the algorithmic definition of Machine Learning with applied, contextual, and socio-institutional approaches, especially in sectors where prediction has direct effects on public, business, or territorial decisions.

According to the observed distribution, theoretical foundations are mainly concentrated in paradigms, with 34 records (43%), and techniques, with 26 records (33%), indicating that the reviewed studies support their approaches in conceptual frameworks of artificial intelligence and machine learning, as well as in quantitative and computational methods to model relationships and make predictions; however, the presence of 17 indeterminate records (22%) reveals a certain degree of conceptual ambiguity in part of the corpus.

This research is pioneering in analyzing definitions of Machine Learning, as it constitutes

the first study to address this type of analysis systematically, which explains the absence of previous studies with which to establish comparisons.

This distribution evidences the need to strengthen theoretical explicitness in future research on Machine Learning and Crime Prediction, in order to improve conceptual traceability, facilitate comparisons among studies, and guide more solid applications in other sectors, regions, and periods of analysis.

4.2.5 RQ5: Which Keywords Present Frequent Co-Occurrence Patterns in Research on Machine Learning and Crime Prediction?

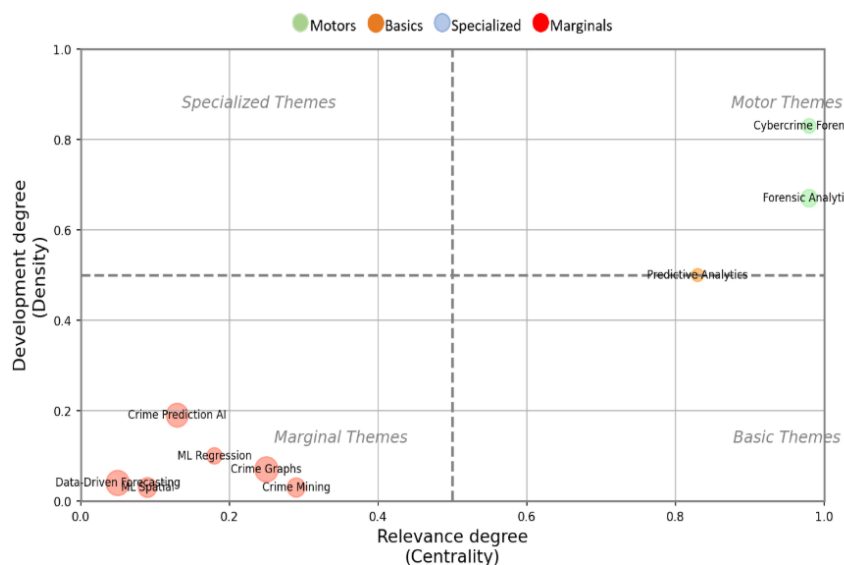
Figure 14, composed of the co-occurrence network and the structural indicators of the keywords, together with Table 12, presents the most frequent association patterns among the terms used in research on Machine Learning and Crime Prediction. Its inclusion is relevant because it makes it possible to identify the dominant thematic nuclei, the intensity of their conceptual links, and the degree of semantic articulation that structures this field of study.

Based on the evidence presented, the terms ml and crime prediction constitute the central conceptual axis of the field, registering the greatest association strength (25.00 in both cases), the highest degree values (0.54 and 0.38, respectively), and the most intense pairwise co-occurrence, corresponding to crime prediction - ml with weight 11, followed by crime prediction - deep learning with weight 8 and deep learning - ml with weight 4; this pattern indicates that the reviewed literature is organized mainly around the direct link between Machine Learning approaches and Crime Prediction, while terms such as data mining, crime analysis, spatial analysis, sentiment analysis, and graph convolutional networks fulfill complementary or emerging functions within a more specialized thematic network.

Barrenechea and Lengua [62] identified that “cybercrime” and “deep learning” were the keywords with the highest co-occurrence, evidencing a relevant thematic relationship between digital crime analysis and the use of advanced Machine Learning techniques. Overall, these findings confirm that these terms constitute important conceptual nuclei for understanding the

Table 12. Total Number of co-occurrences

Keyword 1	Keyword 2	Weight
crime prediction	ml	11
crime prediction	deep learning	8
deep learning	ml	4
crimes	information technology	3
analysis prediction	crimes	2
analysis prediction	information technology	2
crime analysis	data mining	2
crime analysis	ml	2
crime prediction	data mining	2
...	...	...
Total		52

**Fig. 15.** Thematic categories by density and centrality

recent evolution of research on Machine Learning and Crime Prediction.

These results suggest that future research should strengthen the normalization and consistency of keywords, prioritizing central terms such as Machine Learning, crime prediction, and deep learning, and integrating peripheral concepts more explicitly that may broaden the methodological understanding of the field. Likewise, this type of analysis can be replicated in other sectors, geographic regions, and temporal

periods to compare the evolution of thematic nuclei, optimize indexing strategies, and improve the bibliometric visibility of scientific production.

4.2.6 RQ6: Which Thematic Categories Predominate in Research on Machine Learning and Crime Prediction?

Figure 15 and Table 13 present the thematic categories identified from the keywords of the reviewed papers, organized according to two analytical parameters: centrality, located on the

Table 13. Distribution of Thematic Categories by Paper

Theme	Density	Centrality	Qty.	Category
Cybercrime Forensics	0.83	0.98	7	Motor
Forensic Analytics	0.67	0.98	9	Motor
Predictive Analytics	0.50	0.83	9	Basic
Crime Prediction AI	0.19	0.13	22	Marginal
ML Regression	0.10	0.18	13	Marginal
Crime Graphs	0.07	0.25	26	Marginal
Data-Driven Forecasting	0.04	0.05	27	Marginal
Crime Mining	0.03	0.29	15	Marginal
ML Spatial	0.03	0.09	17	Marginal

horizontal axis and associated with the degree of relevance of the theme, and density, located on the vertical axis and related to its level of internal development.

Their inclusion makes it possible to recognize which themes are consolidated, which function as basic axes of the field, and which remain at a marginal stage within research on Machine Learning and Crime Prediction

As shown in the findings, Cybercrime Forensics presents a density of 0.83, centrality of 0.98, and 7 occurrences, while Forensic Analytics reaches a density of 0.67, centrality of 0.98, and 9 occurrences, so both are positioned as motor themes by combining high structural relevance with high conceptual development.

Likewise, Predictive Analytics registers a density of 0.50, centrality of 0.83, and 9 occurrences, which explains its classification as a basic theme due to its high relevance within the field, although with lower internal development than the motor themes. In contrast, Crime Prediction AI (22 occurrences), ML Regression (13), Crime Graphs (26), Data-Driven Forecasting (27), Crime Mining (15), and ML Spatial (17) are located as marginal themes due to their low density and centrality values, indicating that although they are frequent in the keywords, they still present limited theoretical consolidation and low structural articulation within the thematic network.

Garzón Santos and collaborators [75] identified that the terms “artificial intelligence”, “law enforcement”, and “decision support system” present high density and high centrality, evidencing their condition as consolidated and structurally relevant themes within the field; in contrast, “game theory”, “security games”, and “stackelberg games” register low density and centrality, so they represent less developed lines with lower thematic articulation. Complementarily, Kaur and Saini [78] stated that “internet of things”, “IoT”, and “supervised learning” have high density but low centrality, while “data mining”, “crime prediction”, and “big data applications analytics” show high centrality but lower density. Overall, these findings evidence that the predominant thematic categories combine consolidated nuclei, basic themes, and emerging lines, which makes it possible to understand the conceptual evolution of research on Machine Learning and Crime Prediction.

These results indicate that future research should strengthen the articulation between marginal themes and motor themes, especially through studies that integrate predictive models, spatial analysis, data mining, criminal graphs, and forensic approaches, in order to advance toward more robust frameworks for Machine Learning-based Crime Prediction. Likewise, this type of thematic analysis can be replicated in other sectors, geographic regions, and periods of time to evaluate the evolution of research nuclei, guide scientific agendas, and improve the strategy for selecting keywords in future systematic and bibliometric reviews.

The data show that the thematic category with the greatest presence among the analyzed documents was “Data-Driven Forecasting”, with 27 papers, followed by “Crime Graphs” with 26 and “Crime Prediction AI” with 22 papers; however, the “Cybercrime Forensics” category, despite presenting high density and centrality, had reduced participation, with only 7 papers.

In the study by Dakalbab and colleagues [74], “Crime” was the most abundant thematic category with 52 papers (43%), followed by “Spatial” with 19 papers (16%); meanwhile, “Social media” had the lowest presence, with 7 papers (6%).

The results evidence that research on crime prediction tends to concentrate on thematic

categories linked to data analysis, spatial representation, and visualization of criminal patterns; this trend suggests a predominant focus on the use of quantitative technologies and predictive models to address the criminal phenomenon.

5 Conclusions and Future Research

This systematic review analyzed the use of Machine Learning in Crime Prediction based on a final corpus of 60 papers, integrating technical, bibliometric, conceptual, and thematic evidence. Regarding RQ1, the results show that Python is the predominant language for developing Machine Learning-based solutions, due to its library ecosystem, flexibility, and capacity to implement predictive and deep learning models. However, the secondary presence of Java, C++, C#, and PHP indicates that technological selection also depends on integration, architecture, and deployment requirements. Regarding RQ2, applications are mainly concentrated in transportation, mental health, and public safety, evidencing that these models are used in domains where the anticipation of risks, incidents, and behaviors is critical for decision-making. Regarding RQ3, the predominance of publications in Q1 and Q2 journals confirms a documentary base with adequate editorial positioning and scientific impact, although the presence of studies in lower quartiles also requires assessing their methodological and contextual relevance. Regarding RQ4, the literature mostly conceptualizes Machine Learning from a technical perspective, focused on algorithms, statistical models, and computational techniques, revealing methodological strength, but also the need to strengthen theoretical, contextual, and socio-institutional foundations. Regarding RQ6, motor themes, such as Cybercrime Forensics and Forensic Analytics, show high centrality and density, while other frequent themes, such as Crime Prediction AI, Crime Graphs, and Data-Driven Forecasting, still present lower structural consolidation. Overall, the findings evidence that research on Machine Learning and Crime Prediction is moving toward more technical, predictive, and data-oriented approaches;

nevertheless, challenges persist regarding explainability, contextual validation, theoretical integration, and the responsible use of models applied to sensitive decisions in public safety.

Future research should expand the analysis to new periods, regions, and databases to assess the evolution of the field. Likewise, it is necessary to deepen work on explainable models, ethical approaches, and empirical validation in real-world scenarios of safety, mobility, cybersecurity, and risk management. Finally, it is recommended to develop theoretical frameworks that integrate computational, criminological, social, and institutional foundations.

References

1. **Ahmad, R., Nawaz, A., Mustafa, G., Ali, T., Tlija, M., El-Meligy, M. A., Ahmed, Z. (2024).** CHART: Intelligent crime hotspot detection and real-time tracking using machine learning. *Computers, Materials & Continua*, Vol. 81, No. 3, pp. 4171–4194. DOI: 10.32604/cmc.2024.056971.
2. **Alasfoor, E. J., Alshaikh, O., Inuwa-Dutse, I., Khan, S., Parkinson, S. (2025).** Offender Characterization and Prediction: A Case Study of the Kingdom of Bahrain. *IEEE Access*, Vol. 13, pp. 29406–29431. DOI: 10.1109/ACCESS.2025.3531655.
3. **Albahli, S., Alsaqabi, A., Aldhubayi, F., Tayyab Rauf, H., Arif, M., Abed Mohammed, M. (2021).** Predicting the Type of Crime: Intelligence Gathering and Crime Analysis. *Computers, Materials & Continua*, Vol. 66, No. 3, pp. 2317–2341. DOI: 10.32604/cmc.2021.014113.
4. **Alghamdi, J., Al-Dala'in, T. (2023).** Towards Spatio-Temporal Crime Events Prediction. *Multimedia Tools and Applications*, Vol. 83, No. 7, pp. 18721–18737. DOI: 10.1007/s11042-023-16188-x.
5. **Alsubayhin, A., Ramzan, M. S., Alzahrani, B. (2024).** Crime Prediction Model Using Three Classification Techniques: Random Forest, Logistic Regression, and LightGBM. *International Journal of Advanced Computer Science and Applications*, Vol. 15, No. 1, pp.

- 240–251. DOI: 10.14569/IJACSA.2024.0150123.
6. **Zukri, A. Z. M., Sakip, S. R., Masrom, S., Megat, P. R., Zamin, N. (2024).** The Crime Prediction of Criminal Activity Based on Weather Changes towards Quality of Life. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, Vol. 42, No. 1, pp. 130–143. DOI: 10.37934/araset.42.1.130143.
 7. **Armoogum, S., Dewi, D. A., Armoogum, V., Melanie, N., Kurniawan, T. B. (2024).** Unveiling Criminal Activity: A Social Media Mining Approach to Crime Prediction. *Journal of Applied Data Sciences*, Vol. 5, No. 3, pp. 1482–1494. DOI: 10.47738/jads.v5i3.350.
 8. **Aziz, R. M., Hussain, A., Sharma, P., Kumar, P. (2022).** Machine Learning-Based Soft Computing Regression Analysis Approach for Crime Data Prediction. *Karbala International Journal of Modern Science*, Vol. 8, No. 1, pp. 1–19. DOI: 10.33640/2405-609X.3197.
 9. **Aziz, R. M., Hussain, A., Sharma, P. (2023).** Cognizable Crime Rate Prediction and Analysis under Indian Penal Code Using Deep Learning with Novel Optimization Approach. *Multimedia Tools and Applications*, Vol. 83, No. 8, pp. 22663–22700. DOI: 10.1007/s11042-023-16371-0.
 10. **Butt, U. M., Letchmunan, S., Hassan, F. H., Koh, T. W. (2024).** Leveraging Transfer Learning with Deep Learning for Crime Prediction. *PLOS ONE*, Vol. 19, No. 4, e0296486. DOI: 10.1371/journal.pone.0296486.
 11. **Butt, U. M., Letchmunan, S., Hassan, F. H., Koh, T. W. (2022).** Hybrid of Deep Learning and Exponential Smoothing for Enhancing Crime Forecasting Accuracy. *PLOS ONE*, Vol. 17, No. 9, e0274172. DOI: 10.1371/journal.pone.0274172.
 12. **Cichosz, P. (2020).** Urban Crime Risk Prediction Using Point of Interest Data. *ISPRS International Journal of Geo-Information*, Vol. 9, No. 7, 459. DOI: 10.3390/ijgi9070459.
 13. **Dakin, K., Xie, W., Parkinson, S., Khan, S., Monchuk, L., Pease, K. (2020).** Built Environment Attributes and Crime: An Automated Machine Learning Approach. *Crime Science*, Vol. 9, No. 1, 12. DOI: 10.1186/s40163-020-00122-9.
 14. **Dong, Q., Li, Y., Zheng, Z., Wang, X., Li, G. (2022).** ST3DNetCrime: Improved ST-3DNet Model for Crime Prediction at Fine Spatial Temporal Scales. *ISPRS International Journal of Geo-Information*, Vol. 11, No. 10, pp. 529. DOI: 10.3390/ijgi11100529.
 15. **Esquivel, N., Nicolis, O., Peralta, B., Mateu, J. (2020).** Spatio-temporal Prediction of Baltimore Crime Events Using CLSTM Neural Networks. *IEEE Access*, Vol. 8, pp. 209101–209112. DOI: 10.1109/ACCESS.2020.3036715.
 16. **Gelvez Ferreira, J. D., Nieto-Rodríguez, M.-P., Rocha-Ruiz, C.-A. (2022).** Prediciendo el crimen en ciudades intermedias: un modelo de “machine learning” en Bucaramanga, Colombia. *URVIO. Revista Latinoamericana de Estudios de Seguridad*, Vol. 34, pp. 83–98. DOI: 10.17141/urvio.34.2022.5395.
 17. **Gu, H., Sui, J., Chen, P. (2024).** Graph Representation Learning for Street-Level Crime Prediction. *ISPRS International Journal of Geo-Information*, Vol. 13, No. 7, pp. 229. DOI: 10.3390/ijgi13070229.
 18. **Hajela, G., Chawla, M., Rasool, A. (2021).** Crime Hotspot Prediction Based on Dynamic Spatial Analysis. *ETRI Journal*, Vol. 43, No. 6, pp. 1058–1080. DOI: 10.4218/etrij.2020-0220.
 19. **Hardyns, W., Khalfa, R. (2023).** Predicting Crime across Cities and Regions: A Comparative Analysis of Predictive Modelling in Three Belgian Settings. *Applied Spatial Analysis and Policy*, Vol. 16, No. 1, pp. 485–508. DOI: 10.1007/s12061-022-09485-9.
 20. **Houser, T. E., McMillan, A., Dong, B. (2024).** Bridging the Gap between Criminology and Computer Vision: A Multidisciplinary Approach to Curb Gun Violence. *Security Journal*, Vol. 37, No. 4, pp. 1409–1429. DOI: 10.1057/s41284-024-00423-7.
 21. **Khalifa, R., Theinert, N., Hardyns, W. (2025).** Comparing XAI Techniques for Interpreting Short-Term Burglary Predictions at Micro-

- Places. Computational Urban Science, Vol. 5, No. 1, pp. 27. DOI: 10.1007/s43762-025-00185-x.
22. **Kim, D., Jung, S., Jeong, Y. (2021).** Theft Prediction Model Based on Spatial Clustering to Reflect Spatial Characteristics of Adjacent Lands. Sustainability, Vol. 13, No. 14, pp. 7715. DOI: 10.3390/su13147715.
 23. **Kim, G., Cho, Y., Lee, J. H., Lee, G. (2024).** Correlation Analysis Between Urban Environment Features and Crime Occurrence Based on Explainable Artificial Intelligence Techniques. Journal of Asian Architecture and Building Engineering, pp. 1–20. DOI: 10.1080/13467581.2024.2421260.
 24. **Kishore, R. S., Bhaskari, D. L. (2023).** Cyber Crime Detection and Prevention Techniques on Cyber Cased Objects Using SVM and SMOTE. International Journal on Recent and Innovation Trends in Computing and Communication, Vol. 11, No. 8, pp. 259–270. DOI: 10.17762/ijritcc.v11i8.7951.
 25. **Lamari, Y., Freskura, B., Abdessamad, A., Eichberg, S., de Bonviller, S. (2020).** Predicting Spatial Crime Occurrences Through an Efficient Ensemble-Learning Model. ISPRS International Journal of Geo-Information, Vol. 9, No. 11, 645. DOI: 10.3390/ijgi9110645.
 26. **Laqueur, H. S., Copus, R. W. (2024).** An Algorithmic Assessment of Parole Decisions. Journal of Quantitative Criminology, Vol. 40, No. 1, pp. 151–188. DOI: 10.1007/s10940-022-09563-8.
 27. **Matijosaitiene, I., McDowald, A., Juneja, V. (2019).** Predicting Safe Parking Spaces: A Machine Learning Approach to Geospatial Urban and Crime Data. Sustainability, Vol. 11, No. 10, 2848. DOI: 10.3390/su11102848.
 28. **Mihna, F. K. H., Habeeb, M. A., Khaleel, Y. L., Ali, Y. H., Al-Saeedi, L. A. E. (2024).** Using Information Technology for Comprehensive Analysis and Prediction in Forensic Evidence. Mesopotamian Journal of CyberSecurity, Vol. 4, No. 1, pp. 4–16. DOI: 10.58496/MJCS/2024/002.
 29. **Miyano, K., Shinkuma, R., Shiode, N., Shiode, S., Sato, T., Oki, E. (2020).** Multi-UAV Allocation Framework for Predictive Crime Deterrence and Data Acquisition. Internet of Things, Vol. 11, pp. 100205. DOI: 10.1016/j.iot.2020.100205.
 30. **Mohamed Zaidi, M., Avelino Sampedro, G., Almadhor, A., Alsubai, S., Al Hejaili, A., Gregus, M., Abbas, S. (2024).** Suspicious Human Activity Recognition from Surveillance Videos Using Deep Learning. IEEE Access, Vol. 12, pp. 105497–105510. DOI: 10.1109/ACCESS.2024.3436653.
 31. **Mudgal, M., Punj, D., Pillai, A. (2021).** Theoretical and Empirical Analysis of Crime Data. Journal of Web Engineering, Vol. 20, No. 1, pp. 113–128. DOI: 10.13052/jwe1540-9589.2016.
 32. **Niu, X., Elsisy, A., Derzsy, N., Szymanski, B. K. (2019).** Dynamics of Crime Activities in the Network of City Community Areas. Applied Network Science, Vol. 4, No. 1, 127. DOI: 10.1007/s41109-019-0239-8.
 33. **Noor, T. H., Almars, A. M., Alwateer, M., Almaliki, M., Gad, I., Atlam, E.-S. (2022).** SARIMA: A Seasonal Autoregressive Integrated Moving Average Model for Crime Analysis in Saudi Arabia. Electronics, Vol. 11, No. 23, pp. 3986. DOI: 10.3390/electronics11233986.
 34. **Obagbuwa, I. C., Abidoye, A. P. (2021).** South Africa Crime Visualization, Trends Analysis, and Prediction Using Machine Learning Linear Regression Technique. Applied Computational Intelligence and Soft Computing, Vol. 2021, pp. 1–14. DOI: 10.1155/2021/5537902.
 35. **Pouyan, S., Charmi, M., Azarpeyvand, A., Hassanpoor, H. (2023).** Propounding First Artificial Intelligence Approach for Predicting Robbery Behavior Potential in an Indoor Security Camera. IEEE Access, Vol. 11, pp. 60471–60489. DOI: 10.1109/ACCESS.2023.3284472.
 36. **Rayhan, Y., Hashem, T. (2023).** AIST: An Interpretable Attention-Based Deep Learning Model for Crime Prediction. ACM Transactions on Spatial Algorithms and Systems, Vol. 9, No. 2, pp. 1–31. DOI: 10.1145/3582274.

37. **Reier Forradellas, R. F., Nández Alonso, S. L., Jorge-Vazquez, J., Rodriguez, M. L. (2020).** Applied Machine Learning in Social Sciences: Neural Networks and Crime Prediction. *Social Sciences*, Vol. 10, No. 1, pp. 4. DOI: 10.3390/socsci10010004.
38. **Ribeiro, H. V., Lopes, D. D., Pessa, A. A. B., Martins, A. F., da Cunha, B. R., Gonçalves, S., Lenzi, E. K., Hanley, Q. S., Perc, M. (2023).** Deep Learning Criminal Networks. *Chaos, Solitons and Fractals*, Vol. 172, pp. 113579. DOI: 10.1016/j.chaos.2023.113579.
39. **Robertson, D.L., Goodridge, W.S. (2022).** Predicting Density of Serious Crime Incidents Using a Multiple-Input Hidden Markov Maximization a Posteriori Model. *Machine Learning with Applications*, Vol. 7, pp. 100231. DOI: 10.1016/j.mlwa.2021.100231.
40. **Rosés, R., Kadar, C., Malleson, N. (2021).** A Data-Driven Agent-Based Simulation to Predict Crime Patterns in an Urban Environment. *Computers, Environment and Urban Systems*, Vol. 89, pp. 101660. DOI: 10.1016/j.compenvurbsys.2021.101660.
41. **Saeed, R. M., Abdulmohsin, H. A. (2023).** A Study on Predicting Crime Rates through Machine Learning And Data Mining Using Text. *Journal of Intelligent Systems*, Vol. 32, No. 1. DOI: 10.1515/jisys-2022-0223.
42. **Safat, W., Asghar, S., Gillani, S.A. (2021).** Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning Techniques. *IEEE Access*, Vol. 9, pp. 70080–70094. DOI: 10.1109/ACCESS.2021.3078117.
43. **Samia, N., Saha, S., Haque, A. (2024).** Predicting and Mitigating Cyber Threats Through Data Mining and Machine Learning. *Computer Communications*, Vol. 228, pp. 107949. DOI: 10.1016/j.comcom.2024.107949.
44. **Saraiva, M., Matijošaitienė, I., Mishra, S., Amante, A. (2022).** Crime Prediction and Monitoring in Porto, Portugal, Using Machine Learning, Spatial and Text Analytics. *ISPRS International Journal of Geo-Information*, Vol. 11, No. 7, pp. 400. DOI: 10.3390/ijgi11070400.
45. **Sarzaeim, P., Mahmoud, Q. H., Azim, A. (2024).** A Framework for LLM-Assisted Smart Policing System. *IEEE Access*, Vol. 12, pp. 74915–74929. DOI: 10.1109/ACCESS.2024.3404862.
46. **Shan, M., Ye, C., Chen, P., Peng, S. (2025).** Ada GCNLSTM: An Adaptive Urban Crime Spatiotemporal Prediction Model. *Journal of Safety Science and Resilience*, Vol. 6, No. 2, pp. 226–236. DOI: 10.1016/j.jnlssr.2024.11.003.
47. **Stalidis, P., Semertzidis, T., Daras, P. (2021).** Examining Deep Learning Architectures for Crime Classification and Prediction. *Forecasting*, Vol. 3, No. 4, pp. 741–762. DOI: 10.3390/forecast3040046.
48. **Sun, Y., Chen, T., Yin, H. (2023).** Spatial-Temporal Meta-Path Guided Explainable Crime Prediction. *World Wide Web*, Vol. 26, No. 4, pp. 2237–2263. DOI: 10.1007/s11280-023-01137-3.
49. **Tam, S., Tanriover, O. O. (2023).** Multimodal Deep Learning Crime Prediction Using Tweets. *IEEE Access*, Vol. 11, pp. 93204–93214. DOI: 10.1109/ACCESS.2023.3308967.
50. **Tasnim, N., Imam, I. T., Hashem, M. M. A. (2022).** A Novel Multi-Module Approach to Predict Crime Based on Multivariate Spatio-Temporal Data Using Attention and Sequential Fusion Model. *IEEE Access*, Vol. 10, pp. 48009–48030. DOI: 10.1109/ACCESS.2022.3171843.
51. **Umair, A., Sarfraz, M. S., Ahmad, M., Habib, U., Ullah, M. H., Mazzara, M. (2020).** Spatiotemporal Analysis of Web News Archives for Crime Prediction. *Applied Sciences*, Vol. 10, No. 22, pp. 1–16. DOI: 10.3390/app10228220.
52. **Vivek, M., Prathap, B. R. (2023).** Spatio-Temporal Crime Analysis and Forecasting on Twitter Data Using Machine Learning Algorithms. *SN Computer Science*, Vol. 4, No. 4, pp. 383. DOI: 10.1007/s42979-023-01816-y.
53. **Wang, C., Tian, F., Pan, Y. (2023).** Swarm Intelligence Response Methods Based on Urban Crime Event Prediction. *Electronics*,

- Vol. 12, No. 22, 4610. DOI: 10.3390/electronics12224610.
54. **Wu, D.Y., Wang, J. (2024).** Analyzing Urban Crime Through Street View Imagery: Insights from Urban Micro Built Environment and Perceptions. *Urban Science*, Vol. 8, No. 4, 247. DOI: 10.3390/urbansci8040247.
 55. **Wu, J., Abrar, S. M., Awasthi, N., Frias-Martinez, E., Frias-Martinez, V. (2022).** Enhancing Short-Term Crime Prediction with Human Mobility Flows and Deep Learning Architectures. *EPJ Data Science*, Vol. 11, No. 1, 53. DOI: 10.1140/epjds/s13688-022-00366-2.
 56. **Yang, L., Guofan, J., Yixin, Z., Qianze, W., Jian, Z., Alizadehsani, R., Plawiak, P. (2024).** A Reinforcement Learning Approach Combined with Scope Loss Function for Crime Prediction on Twitter. *IEEE Access*, Vol. 12, pp. 149502–149527. DOI: 10.1109/ACCESS.2024.3473296.
 57. **Ye, X., Duan, L., Peng, Q. (2021).** Spatiotemporal Prediction of Theft Risk with Deep Inception-Residual Networks. *Smart Cities*, Vol. 4, No. 1, pp. 204–216. DOI: 10.3390/smartcities4010013.
 58. **Yu, H., Liu, L., Yang, B., Lan, M. (2020).** Crime Prediction with Historical Crime and Movement Data of Potential Offenders Using a Spatio-Temporal Cokriging Method. *ISPRS International Journal of Geo-Information*, Vol. 9, No. 12, pp. 732. DOI: 10.3390/ijgi9120732.
 59. **Yunus, A., Loo, J. (2024).** London Street Crime Analysis and Prediction Using Crowdsourced Dataset. *Journal of Computational Mathematics and Data Science*, Vol. 10, pp. 100089. DOI: 10.1016/j.jcmds.2023.100089.
 60. **Zhang, C., Zhao, X., Cai, M., Wang, D., Cao, L. (2020).** A New Model for Predicting the Attributes of Suspects. *Computer Science and Information Systems*, Vol. 17, No. 3, pp. 705–715. DOI: 10.2298/CSIS200107016Z.
 61. **Alegre-Veliz, R., Gaspar-Ortiz, P., Gamboa-Cruzado, J., Rodríguez Baca, L., Grandez Pizarro, W., Menéndez Mueras, R., Chávez Herrera, C. (2022).** Machine Learning for Feeling Analysis in Twitter Communications: A Case Study in HEYDRUI. *International Journal of Interactive Mobile Technologies*, Vol. 16, No. 24, pp. 126–142. DOI: 10.3991/ijim.v16i24.35493.
 62. **Barrenechea, M. M. M., Lengua, M. A. C. (2025).** A Systematic Literature Review on the Use of Artificial Intelligence for Cybercrime Rate Forecasting. *Bulletin of Electrical Engineering and Informatics*, Vol. 14, No. 3, pp. 2042–2054. DOI: 10.11591/eei.v14i3.9213.
 63. **Cabanillas-Carbonell, M., Rivera, J.S., Muñoz, J.S. (2025).** Artificial Intelligence in Video Surveillance Systems for Suspicious Activity Detection and Incident Response: A Systematic Literature Review. *Advances in Science and Technology Research Journal*, Vol. 19, No. 3, pp. 389–405. DOI: 10.12913/22998624/196795.
 64. **Gamboa-Cruzado, J., Mosqueira-Cerda, T., Torre Camones, A., Quispe Mendoza, R., Navarro Raymundo, A.F., Jiménez García, J., López-Ramírez, B.C. (2025).** Exploring the Influence of Machine Learning in E-Commerce: A Systematic And Bibliometric Review. *Computación y Sistemas*, Vol. 29, No. 2, pp. 925–953. DOI: 10.13053/CyS-29-2-5749.
 65. **Ersoz, F., Ersoz, T., Marcelloni, F., Ruffini, F. (2025).** Artificial Intelligence in Crime Prediction: A Survey with a Focus on Explainability. *IEEE Access*, Vol. 13, pp. 59646–59674. DOI: 10.1109/ACCESS.2025.3553934.
 66. **Esan, O., Chimbo, B. (2025).** Systematic Literature Review on Crime Prediction Using Machine Learning Techniques. *The Indonesian Journal of Computer Science*, Vol. 14, No. 3. DOI: 10.33022/ijcs.v14i3.4881.
 67. **Cárdenas-Quispe, A., Vergaray-Mezarina, R., Gamboa-Cruzado, J. (2021).** Machine learning para la detección de malware en Android: Revisión sistemática de la literatura. *RISTI - Revista Ibérica de Sistemas e Tecnologías de Informação*, E45, pp. 318–331.
 68. **Iqbal, N., Hassan, A., Waheed, T. (2025).** AI Driven Crime Prediction: A Systematic Literature Review. *Journal of Computational*

- Social Science, Vol. 8, No. 2. DOI: 10.1007/s42001-025-00373-z.
69. **Jenga, K., Catal, C., Kar, G. (2023).** Machine learning in Crime Prediction. *Journal of Ambient Intelligence and Humanized Computing*, Vol. 14, No. 3, pp. 2887–2913. DOI: 10.1007/s12652-023-04530-y.
 70. **Gamboa-Cruzado, J., Menendez-Morales, C., Franco del Carpio, C., López-Goycochea, J., Alva Arévalo, A., Ríos Vargas, C. (2023).** Use of Chatbots in E-Commerce: A Comprehensive Systematic Review. *Journal of Theoretical and Applied Information Technology*, Vol. 101, No. 4, pp. 1172–1183.
 71. **Kounadi, O., Ristea, A., Araujo, A., Leitner, M. (2020).** A systematic Review on Spatial Crime Forecasting. *Crime Science*, Vol. 9, No. 1, pp. 1–22. DOI: 10.1186/s40163-020-00116-7.
 72. **Mandalapu, V., Elluri, L., Vyas, P., Roy, N. (2023).** Crime Prediction Using Machine Learning and Deep Learning: A Systematic Review and Future Directions. *IEEE Access*, Vol. 11, pp. 60153–60170. DOI: 10.1109/ACCESS.2023.3286344.
 73. **Rojas Valdivia, J., Gamboa-Cruzado, J., de la Cruz Vélez de Villa, P. (2023).** Systematic Literature Review on Machine Learning and its Impact on Apis Deployment. *Computación y Sistemas*, Vol. 27, No. 4, pp. 1107–1124. DOI: 10.13053/CyS-27-4-4371.
 74. **Dakalbab, F., Abu Talib, M., Abu Waraga, O., Bou Nassif, A., Abbas, S., Nasir, Q. (2022).** Artificial Intelligence, Crime Prediction: A Systematic Literature Review. *Social Sciences & Humanities Open*, Vol. 6, No. 1, pp. 100342. DOI: 10.1016/j.ssaho.2022.100342.
 75. **Garzón Santos, J. L., Ruiz Otálora, J.H., Castro Duarte, J.G., Pérez Pertuz, J.J. (2025).** Importancia de la inteligencia artificial para los cuerpos de policía: un análisis bibliométrico. *Revista Logos, Ciencia & Tecnología*, Vol. 17, No. 1, pp. 119–134. DOI: 10.22335/rict.v17i1.2009.
 76. **Andrade-Mogollon, T., Gamboa-Cruzado, J., Amayo-Gamboa, F. (2025).** Systematic Literature Review of Generative AI and IoT As Key Technologies for Precision Agriculture. *Computación y Sistemas*, Vol. 29, No. 2, pp. 857–882. DOI: 10.13053/CyS-29-2-5738.
 77. **Kitchenham, B., Charters, S. (2007).** Guidelines for Performing Systematic Literature Reviews in Software Engineering (EBSE Technical Report EBSE-2007-01). Software Engineering Group, School of Computer Science and Mathematics, Keele University.
 78. **Kaur, M., Saini, M. (2024).** Role of Artificial Intelligence in the Crime Prediction and Pattern Analysis Studies Published Over the Last Decade: A Scientometric Analysis. *Artificial Intelligence Review*, Vol. 57, No. 8. DOI: 10.1007/s10462-024-10823-1.
 79. **Aparcana-Tasayco, A. J., Gamboa-Cruzado, J. (2022).** Machine Learning for Management in Software-Defined Networks: A Systematic Literature Review. *IEIE Transactions on Smart Processing and Computing*, Vol. 11, No. 6, pp. 400–411. DOI: 10.5573/IEIESPC.2022.11.6.400.

Article received on 06/12/2025; accepted on 16/01/2026.

**Corresponding author is Cesar Jesús Núñez-Prado.*