

## Revista Facultad de Ingeniería

Journal Homepage: <https://revistas.uptc.edu.co/index.php/ingenieria>



# Machine Learning Applied to Gender Violence: A Systematic Mapping Study

Cristian-Camilo Pinto-Muñoz<sup>1</sup>

Jhon-Alex Zuñiga-Samboni<sup>2</sup>

Hugo-Armando Ordoñez-Erazo<sup>3</sup>

**Received:** February 02, 2023

**Accepted:** May 31, 2023

**Published:** June 20, 2023

**Citation:** C.-C. Pinto-Muñoz, J.-A. Zuñiga-Samboni, H.-A. Ordoñez-Erazo, "Machine Learning Applied to Gender Violence: A Systematic Mapping Study," *Revista Facultad de Ingeniería*, vol. 32, no. 64, e15944 2023. <https://doi.org/10.19053/01211129.v32.n64.2023.15944>

## Abstract

Machine Learning (ML) has positioned itself as one of the best tools to address different problems thanks to its data processing capabilities, as well as the different models, algorithms, and predictive factors that help to solve defined problems. Therefore, this article presents a systematic mapping from 2018 to 2023 focused on the application of ML to gender-based violence. The methodology followed for this

<sup>1</sup> Universidad del Cauca (Popayán-Cauca, Colombia). [pintocristian@unicauca.edu.co](mailto:pintocristian@unicauca.edu.co). ORCID: [0009-0006-0545-5700](https://orcid.org/0009-0006-0545-5700)

<sup>2</sup> Universidad del Cauca (Popayán-Cauca, Colombia). [jhonalexsa@unicauca.edu.co](mailto:jhonalexsa@unicauca.edu.co). ORCID: [0009-0007-2887-5908](https://orcid.org/0009-0007-2887-5908)

<sup>3</sup> Ph. D. Universidad del Cauca (Popayán-Cauca, Colombia). [hugordonez@unicauca.edu.co](mailto:hugordonez@unicauca.edu.co). ORCID: [0000-0002-3465-5617](https://orcid.org/0000-0002-3465-5617)



study is based on the definition of elements such as research questions, search strings, bibliographic sources, and inclusion and exclusion criteria. The research results allow us to understand the benefits and challenges of using artificial intelligence, precisely one of its branches, ML, to help combat problems in different areas of society, such as education, health, and violence, among others. It also identifies the countries where ML is being researched and the contexts it is applied to. The study discusses the application of ML to combat gender-based violence. After conducting a literature review, beneficial results were found in the application of artificial intelligence and ML. The results obtained in the different articles showed a predictive capacity and improvements compared to currently used systems. However, despite the positive results, no evidence of the development of an ML model or algorithm applied to gender-based violence in Colombia was found in the review.

**Keywords:** domestic violence; gender-based violence; machine learning; prediction.

### **Machine learning aplicado a la violencia de género: un estudio de mapeo sistemático**

#### **Resumen**

Machine Learning (ML) se ha posicionado como una de las mejores herramientas para abordar diferentes problemáticas gracias a su capacidad de procesamiento de datos y a los diferentes modelos, algoritmos y factor predictivo para ayudar a dar solución a los problemas definidos. Es por ello, que este artículo presenta un mapeo sistemático de los años 2018 a 2023, el cual se orienta en la aplicación de ML enfocado en la violencia de género. La metodología seguida para la realización del estudio parte de la definición de elementos, como preguntas de investigación, cadenas de búsqueda, fuentes bibliográficas y criterios de inclusión y exclusión. Los resultados de la investigación permiten comprender los beneficios y retos que presenta el uso de inteligencia artificial, desde específicamente una de sus ramas, el ML, para ayudar a combatir problemas en diferentes ámbitos de la sociedad, como educación, salud, violencia, entre otros. Además de constatar en qué países se está investigando el ML y en qué contextos es aplicado. El trabajo discute la

aplicación de ML para combatir la violencia de género. Tras realizar el estudio de revisión de la literatura, se encontraron resultados beneficiosos de la aplicación de inteligencia artificial y ML, ya que los resultados obtenidos en los diferentes artículos presentaban capacidad predictiva y mejoras en comparación con los sistemas actualmente usados. Sin embargo, pese a los resultados positivos, no se encontró en la revisión evidencia de desarrollo de un modelo o algoritmo de ML aplicado a la violencia de género en Colombia.

**Palabras clave:** aprendizaje automático; predicción; violencia de género; violencia intrafamiliar.

### **Aprendizado de máquina aplicado à violência de gênero: um estudo de mapeamento sistemático**

#### **Resumo**

O Machine Learning (ML) se posicionou como uma das melhores ferramentas para resolver diferentes problemas, graças à sua capacidade de processamento de dados e aos diferentes modelos, algoritmos e fator preditivo para ajudar a resolver problemas definidos. É por isso que este artigo apresenta um mapeamento sistemático dos anos de 2018 a 2023, que é orientado para a aplicação do ML com foco na violência de gênero. A metodologia seguida para a realização do estudo parte da definição de elementos, como questões de pesquisa, strings de busca, fontes bibliográficas e critérios de inclusão e exclusão. Os resultados da pesquisa permitem compreender os benefícios e desafios apresentados pelo uso da inteligência artificial, especificamente de um de seus ramos, o ML, para auxiliar no combate a problemas em diversas áreas da sociedade, como educação, saúde, violência, entre outros. Além de verificar em quais países o ML está sendo investigado e em quais contextos é aplicado. O trabalho discute a aplicação do ML no combate à violência de gênero. Após a realização do estudo de revisão da literatura, foram encontrados resultados benéficos da aplicação de inteligência artificial e ML, uma vez que os resultados obtidos nos diferentes artigos apresentaram capacidade preditiva e melhorias em comparação com os sistemas utilizados atualmente. No entanto, apesar dos resultados positivos, nenhuma

evidência do desenvolvimento de um modelo ou algoritmo de ML aplicado à violência de gênero na Colômbia foi encontrada na revisão.

**Palavras-chave:** aprendizado de máquina; predição; violência de gênero; violência doméstica.

## I. INTRODUCTION

Currently, one of the major problems present in society is gender-based violence, which is understood as “any act of violence that results or may result, in harm or physical, sexual, or psychological suffering for a woman. It includes threats, coercion, or arbitrary detention, whether it occurs in public or private life” [1]. According to the UN’s (United Nations) report presented in [2], it is estimated that at least 2 million women and female children are at risk of suffering gender-based violence. A significant percentage are indigenous or African descent people living in Colombia’s Pacific coast. This is why it can be said that the danger rates of gender-based violence are projected to be catastrophic. Moreover, according to [3], in Colombia, in recent years, there has been an increase in cases of gender-based violence, particularly sexual and intimate partner violence, which is a problem that needs to be addressed and analyzed.

Considering the abovementioned issues and the relevance of technology, it is noticeable that the use of artificial intelligence (AI) has become more frequent when addressing social problems, given that different methods allow to analyze, observe, and identify key factors that may go unnoticed by the human eye or the traditional methods and systems. One of the main areas of AI is machine learning (ML), which is a type of AI that includes models that can be trained with data to, for example, predict an outcome based on the data used to train the model [4]. Research has shown that the development of different ML models helps face problems in different areas, and there is evidence of its usefulness in several countries, as mentioned in [5][6].

Therefore, ML has been a helpful tool for addressing difficulties in different contexts, such as gender-based violence [7][8]. The literature retrieves some studies where ML is used in gender-based violence. However, there is no evidence of a system or model developed that focuses on the application of ML to tackle gender-based violence in Colombia. This article presents a systematic mapping of ML applied to gender-based violence, and it is divided into three sections. The first section explains the methodology, which presents the elements that will guide the mapping and the criteria for selecting the articles. The second section shows the results obtained after

analyzing the articles selected for the mapping, such as benefits and improvement areas in the application of ML. Finally, the third section presents the conclusions, which expose the information gap on the application of ML to gender-based violence in Colombia and propose future work.

## II. METHODOLOGY

This section explains in detail the process followed in carrying out the systematic mapping of the application of ML to gender-based violence. The methodology is based on [9] and comprises a series of steps explained in detail in the protocol. It includes the questions for the research, the definition of the search strings, the search engines to be used, the libraries to find the articles, and the definition of the inclusion and exclusion criteria.

Moreover, the process described in Figure 1 was followed to complete the systematic mapping. Initially, the research and scope questions were determined. Then, the search of the articles in the databases described in Table 2 took place applying the inclusion and exclusion criteria presented in Table 3. Subsequently, the data were extracted to answer the questions. Finally, the analysis and conclusions were made.

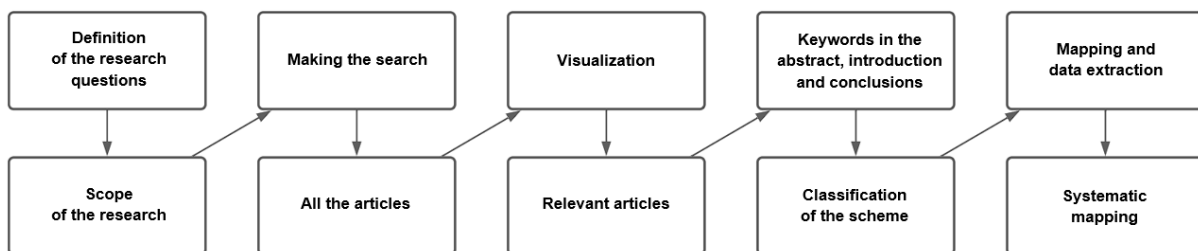


Fig. 1. Process of the systematic mapping.

The general question leading to the objective of the systematic mapping is How are the ML approaches integrated into new and previous gender-based violence prevention and prediction systems?

**Table 1.** Research questions.

Research questions	Motivation
What contributions or innovations does the use of ML applied to different contexts in society involve?	Identifying the application of ML in different contexts of society and knowing if its use has rendered benefits or relevant innovations.
How effective are the ML models in the prediction of gender-based violence compared to the currently used systems?	Identifying the advantages of the tools that apply ML compared to the traditionally used in gender-based violence.
What are the limitations of using ML in the identification or prediction of gender-based violence?	Determining which scenarios cause difficulties, gaps or potential for future work when using ML in gender-based violence.

### A. Protocol

This protocol is based on the proposal of [10]. Each subprocess is described next

**1) Data Sources and Search Strategy.** For this process, the terms in English and Spanish that were used in the search for applications or uses of ML applied to gender-based violence were defined and are presented below: “Machine learning and Gender-Based Violence or domestic violence”, “Machine learning en la violencia intrafamiliar o de género”, “Machine learning violence and abuse”, “Machine learning intimate-partner violence”, “Machine learning family violence”, “Predicción con Inteligencia Artificial y Problemas Sociales”, “Predicción con Machine Learning”, “Inteligencia Artificial en la Violencia contra las mujeres”, “Machine Learning Prediction on social issues”. Table 1 describes the terms used in the search string.

**Table 2.** Search string.

String number	Main terms	Search string
1	Machine learning and Gender-Based Violence or domestic violence	Machine AND Learning AND Gender-Based AND Violence OR domestic AND violence
2	Machine learning en la violencia intrafamiliar o de género	Machine AND learning AND Violencia AND genero OR intrafamiliar
3	Machine learning violence and abuse	Machine AND learning AND Violence AND abuse
4	Machine learning intimate-partner violence	Machine AND learning AND intimate-partner AND violence
5	Machine learning family violence	Machine AND learning AND family AND violence
6	Predicción con Inteligencia Artificial y Problemas Sociales	Predicción AND Inteligencia AND Artificial AND Problemas AND Sociales

## Machine Learning Applied to Gender Violence: A Systematic Mapping Study

String number	Main terms	Search string
7	Predicción con Machine Learning	Predicción AND Machine Learning
8	Inteligencia Artificial en la Violencia contra las mujeres	Inteligencia AND artificial AND violencia AND mujeres
9	Machine Learning Prediction on social issues	Machine AND Learning AND prediction AND social AND issues

Next, Table 3 shows the digital libraries (scientific databases) used to find the selected articles.

**Table 3.** Description of the digital libraries used

Database	Description	URL	Area	Advanced search
Scopus	A comprehensive, expertly selected database of citations and abstracts.	<a href="https://www.scopus.com/">https://www.scopus.com/</a>	Interdisciplinary	Yes
Science Direct	One of the major sources of information for scientific, technical and medical research.	<a href="https://www.sciencedirect.com/">https://www.sciencedirect.com/</a>	Interdisciplinary	Yes
IEEEExplore	A research database to describe articles in areas related to electrical, electronic or software engineering.	<a href="https://ieeexplore.ieee.org/">https://ieeexplore.ieee.org/</a>	Engineering	Yes
Google Scholar	Search engine that allows finding academic, scientific or research articles.	<a href="https://scholar.google.com">https://scholar.google.com</a>	Interdisciplinary	Yes

After defining the search strings and the digital libraries, the inclusion and exclusion criteria used to select the articles were established, as shown in Table 4.

**Table 4.** Inclusion and exclusion criteria.

Type of criterion	Criterion	Type
Period	Articles between 2018 and 2023	Inclusion
Language	Accepted languages were English and Spanish	Inclusion
Type of source	Books, articles, and conferences	Inclusion



Type of criterion	Criterion	Type
Accessibility	Open access articles or articles that can be accessed through agreements of the Universidad del Cauca. Non-accessible	Inclusion Exclusion
Reviewed sections	The alignment of the subject matter in the abstract, introduction, and conclusions was reviewed.	Inclusion
Relevancy for the research questions	Not relevant for one or more research questions	Exclusion

Figure 2 shows the steps followed in selecting the articles, i.e., which sections of the study were analyzed: abstract, introduction, and conclusions. Then, through the keywords observed in each study, it was possible to create the classification scheme and define the subcategories with the different studies according to the topics they address. Including a study in a subcategory of the classification scheme implies ordering and updating the scheme to decide where to place it. Finally, deciding if the study would be included in the systematic mapping.

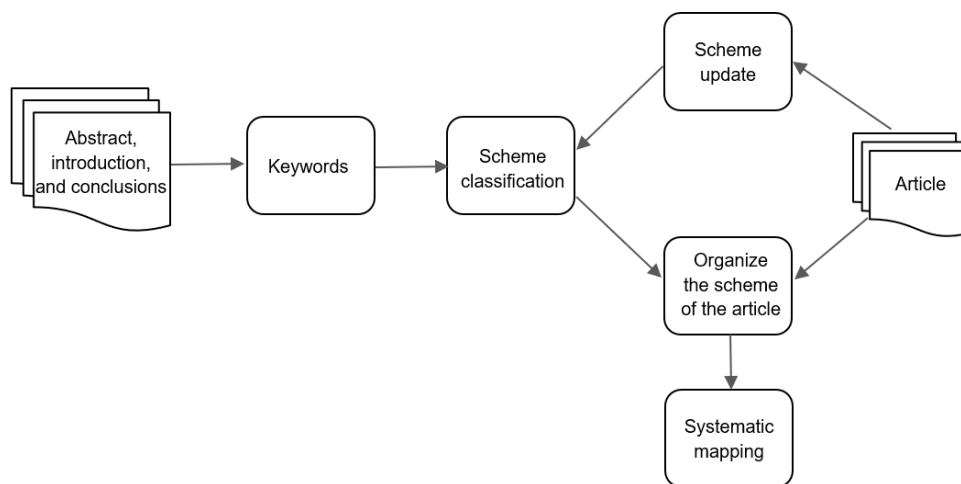


Fig. 2. Classification scheme.

### III. RESULTS

This section presents the results of the search made. It includes a table showing the articles selected from each string and the abstracts of the articles selected for the systematic mapping. Additionally, it presents the answers to the research questions

previously established and a final analysis of the articles. Table 5 shows the result obtained by applying the inclusion and exclusion criteria.

Table 5. Number of results.

Search string	Number of articles found	Selected
1	16,900	5
2	17,300	4
3	14,900	3
4	20,200	6
5	20,100	5
6	16,900	4
7	16,800	11
8	15,900	6
9	18,200	6

In addition to the number of articles found and selected by string, Figure 3 shows the number of publications per year. Most of the selected publications are from 2022, which shows that the topic is currently relevant.

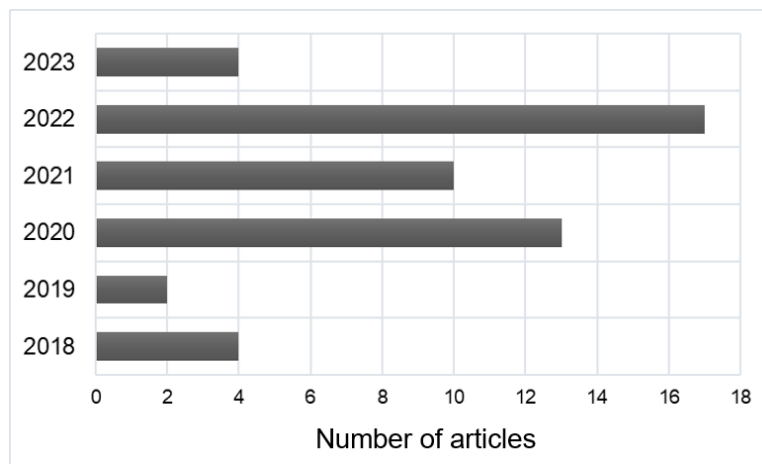


Fig. 3. Publications per year.

Figure 4 shows the main topics addressed by the different studies selected for the mapping. Since the systematic mapping focuses on ML applied to gender-based violence, that topic was the priority when looking for studies on scientific databases. Nevertheless, studies focused on different problems and environments were also selected in order to study and analyze the results of the use of ML in different

contexts, such as ML applied to Colombia, ML focused on violence and different types of violence, and studies focused on the usefulness of ML in the field of health.

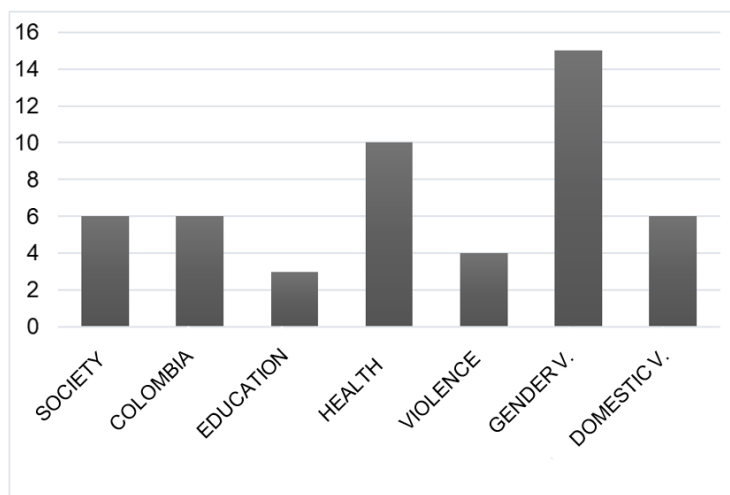


Fig. 4. Topics addressed in the selected studies.

After classifying the studies into categories, their quality was assessed. Then, the criteria presented in Table 6 were defined.

Table 6. Criteria for assessing the quality of the selected studies

Criterion	Description	Score		
		-1	0	1
C1	The study describes in detail the algorithms and methods used.	No	Partially	Yes
C2	The study presents the results clearly and in detail.	No	Partially	Yes
C3	The study is published in a relevant journal.	No	Q4, Q3	Q2, Q1
C4	Other authors have cited the study.	0	1 to 10	10+

Figure 5 shows a chart with the results after assessing the quality of the studies selected for the systematic mapping. It was found that 96% of the studies provided a detailed description of the algorithms and methods used (C1), 90% of the studies showed the results clearly (C2), 54% of the selected studies had been published in

relevant journals (Q1 or Q2), and 22% of the studies had been cited more than 10 times. The majority of the selected studies were found to have been cited 1 to 10 times, accounting for 44%.

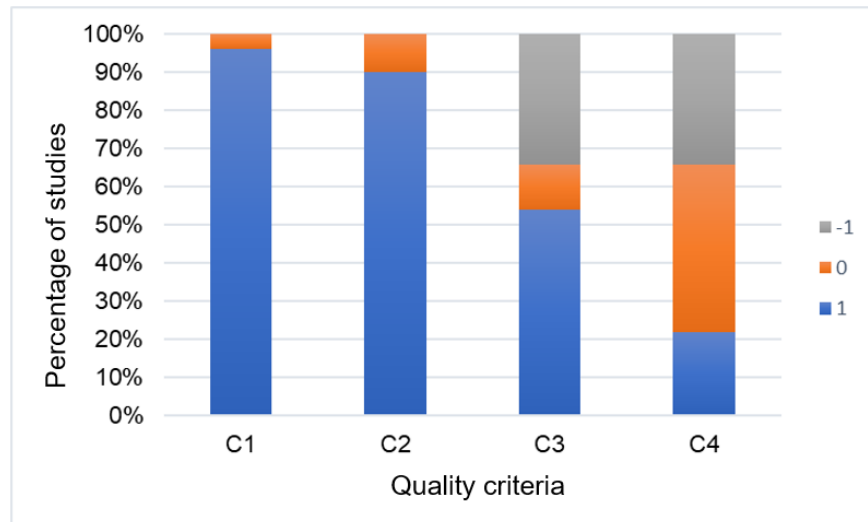


Fig. 5. Distribution of the studies according to the quality criteria assessed.

The following section presents an analysis of the articles selected following the classification criteria.

### A. ML in Society

Wealth distribution, cancer, education, environment, and violence are examples of challenges humanity faces that can be addressed through mathematical tools such as game theory, probability and AI [11]. The latter tool is a source of fear due to the negative narrative by the media and popular culture that see it as a risk that can increase technological unemployment and the creation of lethal autonomous weapons or killer robots. Nonetheless, AI can be used for the common good, for example, for creating tools that help complete the 17 sustainable development goals and their 169 targets proposed in the 2030 agenda [12].

Among the many fields of AI, ML is of great help in the social sciences because it is applied to create models that provide predictions that help make political decisions, develop theories [13] or even predict socioeconomic vulnerabilities in rural areas considering factors such as remoteness, geographical conditions or socioeconomic

dependence on primary sectors, which helps to have better planning for rural development [14].

Rural anomalies have also been studied with ML. In [15], ML and other technologies, such as big data, were used to predict and detect urban anomalies divided into four categories. The concern with anomalies is that they are critical events that can potentially cause the loss of one or multiple lives. Finally, favorable results have also been seen in the use of ML as a tool to fight climate change. In [16], the importance of forests is emphasized. The study applied ML to estimate and predict carbon reservoirs in a tropical forest in southeastern Brazil, showing satisfactory performance.

### ***B. Application and Use of ML in Colombia***

In Colombia, the application of tools such as AI and ML has been focused on areas such as agriculture to provide a better understanding of the patterns and dynamics of nature and predict physical variables of the soil, which helps to make better decisions for the crops after analyzing them [17]. It is also possible to identify essential factors in crop yields, as can be seen in [18], where ML algorithms were used to predict variables related to the productivity of cocoa crops in Santander, which provides farmers with an opportunity to increase their income.

Another area in which AI is applied in Colombia is education, where it has been used to predict academic performance in higher education students based on the most influential factors (educational, family, socioeconomic, habits, among others), which are used then to classify new students so that they can have greater support according to the prediction of their performance [19].

ML has also been used in Colombia to predict theft tendencies. In [20], the behavior of future criminal acts in three main cities (Bogotá, Medellín, and Cali) is predicted using ML based on adjusted support vector machines for regression (SRV). Similarly, ML was used to predict violent homicide tendencies. In [21], a supervised ML model based on random forest was used. The data obtained by the model were close to the ideal data and forecasted a decrease in violent homicide rates in the future. However, studies have not only been conducted in major cities in Colombia.

In [22], ML was used to prevent crime in Bucaramanga, an intermediate city, in which a TF-IDF text vectorization model was applied, and it was concluded that its application to predict crime in intermediate cities is a useful tool but that it may be limited in case there is little recorded crime information.

### ***C. ML Applied to Education***

One of the contexts in which there is also ample evidence and previous studies on the use and application of AI and ML is the educational environment. Multiple articles have addressed topics related to education, such as student desertion.

In [23], an ML algorithm was used to predict university desertion in Perú, obtaining a 91% success rate in the prediction. There are also articles on the same topic in Colombia. In [24], data mining techniques were used to identify the factors or variables that influence student desertion the most; the effectiveness in identifying the most influential variables was greater than 80%. However, student desertion is not the only application of AI in education. In [25], ML methods were used in the prediction of psychosocial risks in teachers at Colombian public schools. These factors can have effects on the health and security of the teachers. The results of the article present an effectiveness of 93% in the prediction of psychosocial risks. The areas of application of ML in the educational field are promising since they offer educational institutions a great opportunity to address problems or phenomena appropriately through the results obtained from the application of ML.

### ***D. ML Applied to Health***

In the health field, the study and use of ML has great innovation potential, considering that most of the studies are based on traditional surveys and do not take advantage of the predictive power of ML to address the social determinants of health [26], even though its use has been increasing. For example, an algorithm based on ML was created to predict the risk of having acute coronary syndrome [27]. Moreover, in Ecuador, ML models were used to optimize the time required for a nutritionist to calculate the BMI of their patients [28] since the inhabitants of a specific area were having health issues. ML was also used to predict breast cancer by

creating a neural network with biomarkers as the information source to decrease the time when cancer can be ruled out [29]. This study had promising results with an accuracy of 82.76% and demonstrated that it could be used as a complementary tool for predicting breast cancer.

Predictive models can be used to predict the risk of mental illnesses such as depression in older adults based on predictors such as age, physical health, and cognitive function [30]. Similarly, they can be used in psychosocial problems, as in [31], where the emphasis is placed on the importance of decision-making in the case of urban bus drivers, who have one of the professions with the highest risk and stress load. Thus, the application of ML for predicting occupational psychosocial risk for urban transport drivers is presented as a great tool to address this problem, thanks to the good performance presented by the model.

Models have also been applied in higher-risk areas, such as the use of electrocardiograms (ECGs) to predict mortality risk in patients presenting to the emergency room. The performance of ECG-based algorithms was good to excellent, with an area under the receiver operating characteristic curve of 0.843 [32].

The COVID-19 pandemic brought many problems to the world, causing irreparable damage. Currently, there are several variants, each with different symptoms, which, if not treated in time, can lead to the death of patients. Therefore, a predictive model that uses ML to identify the presence of the disease has been created to help health specialists to provide a fast, efficient, and reliable service [33].

Another relevant area during the COVID-19 pandemic was mental health issues and the impact on the quality of life due to the confinement established almost in every country. In [34], ML is used to predict the relationships between the mental health of people during COVID-19 and demographic and social factors, showing as a result that the most vulnerable people at the time of confinement were those with a previous diagnosis of mental illness. However, the prevalence of psychiatric problems also increased in the rest of the population. Furthermore, [35] used ML to identify child abuse and intimate partner violence predictors during COVID-19. The ML method used was random forest and SHAP values fed with data provided by 380 people with an average age of 36 years. The risk factors associated with COVID-19

were predicted, such as parental psychological distress (anxiety, depression). There was also an increase in the physical punishment of children and verbal fights between couples.

### **E. ML as a Tool to Address Violence**

Models have been proposed to predict the escalation and de-escalation of violence. For example, in [36], the history of the conflict and the characteristics of the text collected from more than 3.5 million newspaper articles were used as sources of information. With the data mentioned above, the ML model was implemented on the topics covered in each of the newspaper articles, obtaining as a result that this model is more useful in countries with a long history of war, such as Venezuela, where it can predict whether the conflict will increase or decrease. However, it can also be used to warn of the outbreak of unexpected conflicts in peaceful countries.

On the other hand, firearm violence is a common problem in many countries, such as the United States, where it is one of the leading causes of death and injuries. In [37], ML was used to identify factors associated with firearm violence in urban areas to provide relevant information that can be used to analyze violence and health. The study obtained 18 predictive co-variables that explain 77.8% of the variation in the rate of violence with firearms, where associated factors such as isolation indexes, schooling rates, and wealth and poverty indicators, among others, can be found.

Likewise, the application of ML can be a great tool in violence-oriented contexts, an area that comprises different types of situations that affect life in society. Thanks to the advancement of technology, it is possible today to have records and detailed information about events that occur over specific periods. All this stored information is very useful for tools that can process and use that enormous amount of data. [38] compares the use of different ML algorithms to predict crime. The prediction of this problem is one of the most promising objectives since developing a tool that predicts when a crime will occur would allow law enforcement organizations to act efficiently and reduce crime levels. [39] reviews the theory of the techniques and assumptions regarding crime prediction employing ML tools.



## ***F. Use of ML in the Context of Gender-Based Violence***

Many resources are put into fighting gender-based violence, and ML can be a great tool in addressing this problem. [40] discusses the role of technology in the prevention of gender-based violence from the perspective of women and attempts to demonstrate how technology can be a crucial factor, allowing better coverage and attention to cases of gender-based violence and its prevention. One way of preventing gender-based violence is detecting dangerous situations. [41] presents the results of a pilot test where neurotransmitters are used to detect emotions. Furthermore, [42] applies artificial neural networks to violence against women by their intimate partners.

Some studies address gender-based violence by considering specific elements, such as [43], where an ML model was used to find patterns in the cases of fatal gender-based violence victims that lived with their partners. Similarly, in [44], 4 different ML algorithms were used and compared to predict the number of reports of gender-based violence in Spain. The dataset used was prepared from data and characteristics related to gender-based violence collected directly from the official INE (National Statistics Institute) sources. Different feature selection techniques were applied to choose the most critical and influential features, resulting in faster training for the ML algorithms. After training, it was possible to predict reports of gender-based violence with acceptable accuracy and determine the predictive algorithm with the best results, in this case, random forest.

Likewise, studies are addressing gender-based violence from different approaches, such as social media. In [45], they sought to identify factors that influence a victim of intimate partner violence to decide to stay or leave the abusive relationship. The study applied ML to analyze the human notation of tweets and hashtags in the Twitter application. As a result, a classifier of reasons with 78% accuracy was obtained.

In [6], ML models were trained to evaluate patients for IPV (Intimate partner violence) and injuries. These models were trained on radiology reports, and it was detected that victims of intimate partner violence have a higher risk of presenting health problems. Moreover, in [46], using supervised learning models, the relationship

between IPV and suicides was found, showing that it is possible to detect cases of intimate partner violence in suicides with an accuracy of 78%. However, the vague language presented in the reports is a limitation.

IPV was also the subject of study in [47], where 4 different ML algorithms were applied to predict factors associated with the risk of experiencing IPV in South Africa. The models were trained with data from surveys of more than 1500 women. After their training phase, the algorithms confirmed previous studies conducted on factors that affect intimate partner violence and identified factors or unique combinations of factors that constitute a greater or lesser risk of experiencing intimate partner violence. Finally, after comparing the algorithms used, the algorithm with better prediction performance and greater sensitivity was the DT (decision tree) algorithm. The problem of gender-based violence and the study of tools to help tackle it have also been the subject of study in Spain, where there is an official system called VioGen, which seeks to monitor and control gender-based violence. In [48],[48] ML was used to predict the recidivism of an aggressor. The data fed to the models were obtained through VioGen, and the results after applying ML show an improvement in the performance of up to 25% compared to the assessment system currently used by VioGen. Additionally, a study that also demonstrated the effectiveness of using ML versus the current risk assessment model was [7], where it was observed that a large amount of data and responses allows an ML-based model to extract subtle information that would be difficult to obtain by other classical means.

Like Spain, another country deeply affected by gender violence is India, where 6% of women have experienced sexual violence, which is 1.8% higher than the world average, and despite various initiatives to reduce the percentage, no positive results have been seen. For this reason, in [49], ML was used to explore a range of social, demographic, and health factors provided in a large national survey of 66,013 women aged 15-49 years who were ever married to determine which of these factors are associated with marital sexual violence (MSV). Similarly, in [50], a study was conducted applying regularized regression models to identify factors associated with non-marital sexual violence (NMSV), having as a source of information the national family health survey conducted between 2015 and 2016, where having previously

suffered this type of violence was identified as the main risk factor, followed by geography, sexual behavior, and low knowledge about sexual and reproductive health (SRH). Moreover, an important factor to consider is the low number of women that seek help. Reports say only 14% of women who suffered intimate partner violence sought help. Therefore, [51] used ML to identify potential factors associated with the decision to seek help. The regression ML model used allowed to identify key factors that influence the decision to seek help, formally or not.

### ***G. Application of ML in Domestic Violence***

Domestic violence is a type of violence that negatively affects society and the core of a family. Children and adolescents are the main victims of domestic violence. Thus, [52] proposes the use of ML techniques to predict the involvement of adolescents in family conflicts.

Due to phenomena such as the COVID-19 pandemic, there is an increase in social problems given the high levels of anxiety and stress; one of these problems is domestic violence. Therefore, technological tools such as ML are being developed to propose solutions to tackle these issues.

In social media, there is an immense amount of data stored that can be of great use for the application and use of ML. In [53], an ML approach is used to analyze more than one million tweets. Although statistical studies have already been conducted, the approach with the use of ML allows for studying, forecasting, and predicting the impacts on domestic violence based on factors such as the restrictions during COVID-19 [54].

Likewise, in Bangladesh during the COVID-19 outbreak in the summer of 2020, factors such as preventive measures and restrictions led to an increase in social problems, one of them being domestic violence against women and children. The use of different ML algorithms to predict domestic violence was proposed. The methods applied were random forest, logistic regression, and Naive Bayes. Virtual surveys were applied during the COVID-19 outbreak to collect the data for the model, obtaining responses from 511 families. The results obtained show ML can predict

domestic violence with acceptable accuracy. The logarithm with the highest prediction was random forest [55].

In Argentina, [56] presents a predictive model that defines the level of risk of people reporting domestic or gender-based violence between 2017 and 2020. In this model, different variables were used (number and types of violence reported, gender, age of the victim, etc.) to obtain systematic predictions with acceptable assertiveness; AUC=0.82 (area under the curve), and risk factors were observed. For example, among the people who report gender-based violence, 76% are women, and 27% are people between 0 and 14 years of age, showing that the primary victims of domestic violence are women and children. It was also possible to predict that 7% of the reports have a very high risk, 32% have a high risk, and 61% have a medium, moderate, low or no risk.

In the United Kingdom, [8] identified that the current system used to provide adequate assistance to victims of domestic abuse, known as DASH (Domestic Abuse, Stalking and Honour Based Violence), is not useful for identifying the most vulnerable victims. Therefore, the study evaluates several ML algorithms to determine if they can identify the most vulnerable victims and which model performed the best. The dataset used to develop and evaluate the model has 350,000 domestic violence cases, divided into IPV and non-IPV. The results showed that it is possible to determine victims at risk better using the ML model compared to the DASH tool. The best-performing model was the logistic regression model with an elastic net.

## **H. Answers to the Research Questions**

Following an analysis of the selected articles, the answers to the research questions defined in Table 1 are presented below.

**1) What Contributions or Innovations does the use of ML Applied to Different Contexts in Society Involve?** ML has been useful in several contexts. In the society in general, it can be applied to politics, socioeconomics, environment, wealth, or the development of theories, due to its capability to predict multiple scenarios [11],[12], [13],[14]. On a more specific level, it has been used in health, where it is very

innovative since it can deliver traditional surveys and predict the risk levels of different diseases such as cancer, angina, or mental illnesses [26],[27],[29],[30],[32]. It has also been useful in dealing with unforeseen health issues that have led to further problems, such as the COVID-19 pandemic and its repercussions on people's lives. ML was an efficient tool in the identification and prediction of COVID-19 cases [33], the determination of other factors such as child abuse and violence against women during the pandemic [35], and the prediction of the mental health of the people during the pandemic [34].

ML has also been used in the educational field to determine the student desertion rate, the psychosocial risks in teachers, and the classification of students according to different variables that allow teaching according to their abilities and knowledge [19],[23],[24],[25].

In Colombia, studies are addressing the application of artificial intelligence to different fields, such as agriculture, to predict crop productivity [17],[18], education [19], and crime [20],[21],[22].

ML is presented as a great opportunity since the models and algorithms allow predicting or estimating behaviors from the processing of existing information, allowing to address various issues with a predictive approach and giving the possibility to the entities in charge of making the most appropriate decisions regarding the needs or problems they try to solve [15],[16],[28],[31].

In more recent years, ML has also been very useful in facilitating decision-making in challenging circumstances or environments in which no previous experience has been available, such as in the confinements due to COVID-19 and their consequences [33],[34],[35]. It has also been a useful tool in high-risk problems, such as violence, allowing one to predict its behavior based on different factors. Efforts have also been made to specifically study negative behaviors in society, such as crimes, thefts, and homicides, among others [20],[21],[22],[36],[37],[38],[39].

The studies that were found and selected allow us to notice an approach to violence, with an emphasis on domestic violence and gender-based violence, through ML. The number of publications and studies evidence the joint efforts to try and solve or help with these issues. Regarding domestic violence, some articles try to determine

different factors, such as level of risk [56], level of adequate assistance [8], and prediction of domestic violence according to specific factors [55], among others [52],[53],[54].

In the case of gender-based violence, multiple articles propose different objectives, such as proposing new systems or improving current systems [5], comparing different algorithms to find the one with the best performance for a given problem [8], determining situations of danger [41], predicting the number of reports of gender violence [44], minimizing the recidivism of the aggressor [48], determining factors associated with sexual violence [49], and determining the appropriate level of protection for a victim [5]. These are some of the objectives that can be established when addressing gender-based violence with an ML approach [6],[7],[40],[42],[43],[45],[46],[47],[49],[50],[51].

## **2) How Effective are the ML Models in the Prediction of Gender-Based Violence Compared to the Currently Used Systems?**

Spain has the viogen system, which has been in operation since 2007 and has more than 600,000 cases of gender-based violence registered by 2020. Due to the large amount of data registered by viogen and the processing capacity of ML, it was possible to design and test ML models that were trained with a dataset of over 40,000 reports. Subsequently, it was possible to observe an improvement by ML models in the prediction of violence cases and factors, such as the level of police protection, where a 30% improvement was seen compared to other models and systems, including the currently used system, viogen [5]. Other studies have also shown a 25% improvement in the accuracy of gender-based violence predictions using ML algorithms with respect to viogen. This is due to the large source of information available in this country, allowing the extraction of subtle information through ML-based algorithms that could not be reflected using conventional statistical methods [7],[48].

Another place where it was possible to demonstrate the effectiveness of ML compared to conventional tools was in the United Kingdom, where they compared the DASH system with a predictive model using logistic regression with an elastic net which had data from a police department responsible for a metropolitan area in the United Kingdom that included 350,000 incidents of domestic abuse. In this study,

an improvement of .748 Area Under the Curve (AUC) was seen in the prediction of intimate partner violence and of .763 AUC for other forms of domestic abuse. In addition, it was shown that the questions presented in the DASH system were not very useful when predicting the risks suffered by the most vulnerable populations [8].

### **3) What are the Limitations of Using ML in the Identification or Prediction of Gender-Based Violence?**

- It has been shown that having a smaller source of information will increase the error rate and its dispersion, causing greater prediction variation [44].
- The lack of detailed language in the reports presented limits the use of ML tools by not having sufficiently clear information [46].
- Spain has a current system that seeks to minimize recidivism in cases of intimate partner violence. It was noticed that the models adjusted with fewer variables performed worse [7].
- Data collected in violence cases are affected by different biases, such as recall bias or social desirability bias [47],[49].
- It is difficult to predict certain factors, such as risk. A predictive model was used to predict the level of risk and thus allow the police to provide a level of security to the victim. It is a complex task to determine whether the victim was not in danger later because of the security provided or because the prediction of the level of risk failed [5].
- There may be low accuracy in violence prediction if low fidelity data is used as a source of information, e.g., in [45], data is collected from the social media Twitter, which may contain jokes, spam or messages posted by bots.

## **IV. ANALYSIS**

Figure 6 shows the heat map corresponding to the countries where the studies of the papers selected for the systematic mapping of the literature were carried out. However, some studies were excluded, given that they did not have a specific



country where the study took place, such as [53], where the source of information were tweets on a global scale.

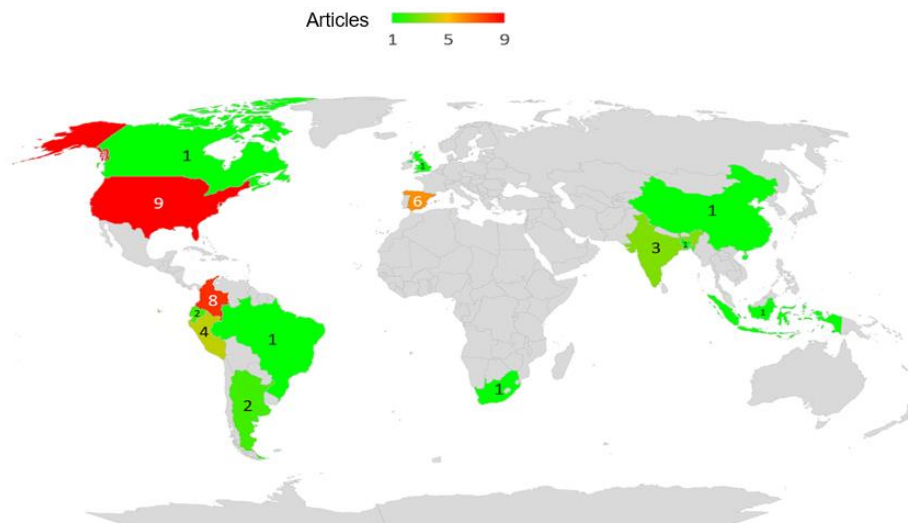


Fig. 6. Heat map of the studied countries.

Figure 6 shows that the most studied countries in the models that use ML are the United States, Colombia, and Spain. In the United States and Spain, studies were found that propose or apply ML models to prevent gender-based violence. In Spain, several studies compare ML-based models with the tool currently used (VioGen), showing, in all cases, an improvement in the accuracy of violence prediction using ML. On the other hand, in Colombia, the studies are aimed at agriculture or violence in general, but no evidence was found of tools applied to prevent gender-based violence.

Employing the VOSviewer software, a keyword correlation map was made, as shown in Figure 7. This map was made on one of the established search strings in the Scopus bibliographic database. Word correlation analysis helped analyze the correlation and co-occurrence of words in the different articles. Four clusters were obtained, represented by four colors: blue, red, yellow, and green. The interpretation is that the terms that share a group have a higher correlation. The term with the highest concurrence was “machine learning”, i.e., it is the word most mentioned in the articles. Other interesting keywords that had a high concurrence and allowed finding influential terms in the articles were “gender-based violence”, “risk factor”,





## Machine Learning Applied to Gender Violence: A Systematic Mapping Study

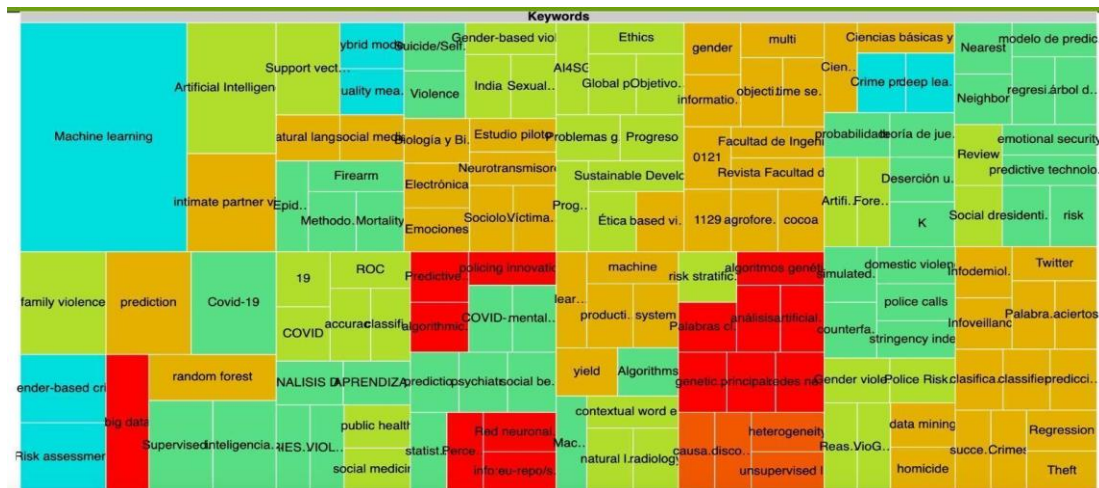


Fig. 8. Keywords treemap.

With the information shown in the systematic mapping, it was possible to notice that many studies are concentrated on nations such as India and Spain on gender violence. On the other hand, in Colombia, some studies and articles prove and promote the use of ML in contexts such as suicides, crimes and others, but no evidence was found of tools that apply ML techniques for the prevention of cases of gender-based violence, despite the catastrophic projection in this type of violence presented by the UN, which shows an estimate of two million women and children at risk [2]. Therefore, the development of ML models and tools would be very useful to support decision-making by the state in the implementation of policies that help prevent and minimize the risk of gender violence.

## V. CONCLUSIONS

The selected studies highlight the role that ML plays in the resolution of problems and difficulties that affect today's society, occupying a role where its objective is to serve as a tool to create innovative systems or improve current systems that tackle problems in multiple contexts. That is why the selected publications present the benefits of the use of ML in different scenarios, such as agriculture, health, education, crime, and violence, among others.

Tools based on ML oriented to violence face one of the main difficulties of current prevention models since they are limited to presenting statistical data on cases that

have occurred in a specific period but do not show data for the future, which may make it impossible to carry out prevention measures on outbreaks of violence that occur after the data presented. On the contrary, with a system or a tool based on ML algorithms, such information can be used to make estimates and predictions, being useful to predict possible cases of violence, determine the most influential factors from the data collected, study areas more prone to violence, support prevention and aid campaigns, and fight the problem of violence in a better way.

Due to the information gap found regarding the use of ML for the prevention of gender-based violence in Colombia, as future work, we expect to conduct research on this topic in order to develop a dashboard-type web tool that implements an ML model based on supervised algorithms that will serve as a contribution to the national public policy support.

#### **AUTHORS' CONTRIBUTION**

**Cristian-Camilo Pinto-Muñoz:** Conceptualization, Methodology, Writing-original draft, Supervision.

**Jhon-Alex Zuñiga-Samboni:** Conceptualization, Methodology, Writing-original draft.

**Hugo-Armando Ordoñez-Erao:** Supervision, writing-review and editing.

#### **ACKNOWLEDGMENTS**

The authors express their gratitude to the Universidad del Cauca for their collaboration and assistance during the creation of the article "Machine Learning Applied to Gender Violence: A Systematic Mapping Study".

#### **REFERENCES**

- [1] UN Women, Fourth World Conference on Women, 1995. <https://www.un.org/womenwatch/daw/beijing/platform/violence.htm#diagnosis>
- [2] UN Women, *Las mujeres en Colombia*, 2023. <https://colombia.unwomen.org/es/ONU-mujeres-en-colombia/las-mujeres-en-colombia>
- [3] ReliefWeb, *Colombia: Situación de la Violencia Basada en Género (VBG) - Comparativo 2020 – 2021*, 2022. <https://reliefweb.int/report/colombia/colombia-situacion-de-la-violencia-basada-en-genero-vbg-comparativo-2020-2021-abril>
- [4] IBM, *¿Qué es Machine Learning?*, 2023. <https://www.ibm.com/mx-es/analytics/machine-learning>

## Machine Learning Applied to Gender Violence: A Systematic Mapping Study

- [5] Á. González-Prieto, A. Brú, J. C. Nuño, J. L. González-Álvarez, "Machine learning for risk assessment in gender-based crime," *Arxiv*, Jun. 2021. <https://doi.org/10.48550/arxiv.2106.11847>
- [6] I. Y. Chen, E. Alsentzer, H. Park, R. Thomas, B. Gosangi, R. Gujrathi, B. Khurana, "Intimate Partner Violence and Injury Prediction from Radiology Reports," *Biocomputing*, vol. 26, pp. 55–66, 2021. [https://doi.org/10.1142/9789811232701\\_0006](https://doi.org/10.1142/9789811232701_0006)
- [7] L. Quijano-Sánchez, F. Liberatore, G. Rodríguez-Lorenzo, R. E. Lillo, J. L. González-Álvarez, "A twist in Intimate Partner Violence Risk Assessment Tools: Gauging the contribution of exogenous and historical variables," *Knowledge-Based Systems*, vol. 234, e107586, 2021. <https://doi.org/10.1016/j.knosys.2021.107586>
- [8] E. Turner, G. Brown, J. Medina Ariza, "Predicting Domestic Abuse (Fairly) and Police Risk Assessment," *Psychosocial Intervention*, vol. 31, no. 3, pp. 145-157, 2022. <https://doi.org/10.5093/pi2022a11>
- [9] B. A. Kitchenham, D. Budgen, O. Pearl Brereton, "Using mapping studies as the basis for further research - A participant-observer case study," *Information and Software Technology*, vol. 53, no. 6, pp. 638–651, 2011. <https://doi.org/10.1016/j.infsof.2010.12.011>
- [10] A. Carrera-Rivera, F. Larrinaga, G. Lasa, "Context-awareness for the design of Smart-product service systems: Literature review," *Computers in Industry*, vol. 142, e103730, 2022. <https://doi.org/10.1016/j.compind.2022.103730>
- [11] F. Isa, "Propuestas de modelos originales en inteligencia artificial, Teoría de juegos y probabilidad para: distribución de riquezas, medio ambiente, alimentos, cáncer, violencia y educación," *South Florida Journal of Development*, vol. 3, no. 5, pp. 6171–6191, 2022. <https://doi.org/10.46932/sfjdv3n5-053>
- [12] A. Monasterio Astobiza, "Inteligencia Artificial para el bien común (AI4SG): IA y los Objetivos de Desarrollo Sostenible," *Arbor*, vol. 197, no. 802, e629, 2021. <https://doi.org/10.3989/arbor.2021.802007>
- [13] M. Molina, F. Garip, "Machine Learning for Sociology," *Annual Review of Sociology*, vol. 45, pp. 27–45, Jul. 2019. <https://doi.org/10.1146/annurev-soc-073117-041106>
- [14] D. Yuliawan, D. Budiman Hakim, B. Juanda, A. Fauzi, "Classification and prediction of rural socio-economic vulnerability (IRSV) integrated with social-ecological system (SES)," *Decision Science Letters*, vol. 11, pp. 223-234, 2022. <https://doi.org/10.5267/j.dsl.2022.4.001>
- [15] M. Zhang, T. Li, Y. Yu, Y. Li, P. Hui, Y. Zheng, "Urban Anomaly Analytics: Description, Detection, and Prediction," *IEEE Trans Big Data*, vol. 8, pp. 809-826, 2022. <https://doi.org/10.1109/tbdata.2020.2991008>
- [16] D. Dantas, M. de Castro Nunes Santos Terra, L. P. Baldissera Schorr, N. Calegarioa, "Machine learning for carbon stock prediction in a tropical forest in Southeastern Brazil," *Bosque (Valdivia)*, vol. 42, no. 1, pp. 131–140, 2021. <https://doi.org/10.4067/S0717-92002021000100131>
- [17] H. S. Garcia Cañon, "Implementación de técnicas de machine learning para la predicción de variables meteorológicas y del suelo que afectan la agricultura," Grade Thesis, Universidad de Los Andes, Colombia, 2019. <https://repositorio.uniandes.edu.co/bitstream/handle/1992/45458/u827540.pdf>
- [18] H. Lamos-Díaz, D. E. Puentes-Garzón, D. A. Zarate-Caicedo, "Comparison Between Machine Learning Models for Yield Forecast in Cocoa Crops in Santander, Colombia," *Revista Facultad de Ingeniería*, vol. 29, no. 54, e10853, 2020. <https://doi.org/10.19053/01211129.V29.N54.2020.10853>
- [19] O. D. Castrillón, W. Sarache, S. Ruiz-Herrera, "Predicción del rendimiento académico por medio de técnicas de inteligencia artificial," *Formación universitaria*, vol. 13, no. 1, pp. 93-102, 2020. <https://doi.org/10.4067/S0718-50062020000100093>
- [20] H. Ordóñez, C. Cobos, V. Bucheli, "Machine learning model for predicting theft trends in Colombia," *Revista Iberica de Sistemas e Tecnologias de Informacao*, vol. 2020, no. E29, pp. 494-506, 2020.
- [21] H.-A. Ordoñez-Eraso, C.-J. Pardo-Calvache, C.-A. Cobos-Lozada, "Detección de tendencias de homicidios en Colombia usando Machine Learning," *Revista Facultad de Ingeniería*, vol. 29, no. 54, e11740, 2020. <https://doi.org/10.19053/01211129.v29.n54.2020.11740>
- [22] J. D. Gelvez Ferreira, M. P. Nieto Rodríguez, C. A. Rocha Ruiz, "Prediciendo el crimen en ciudades intermedias: un modelo de 'machine learning' en Bucaramanga, Colombia," *Urvio. Revista Latinoamericana de Estudios de Seguridad*, no. 34, pp. 83-98, 2022. <https://doi.org/10.17141/urvio.34.2022.5395>
- [23] J. E. Valero Cajahuanca, Á. F. Navarro Raymundo, A. C. Larios Franco, J. D. Julca Flores, "Deserción universitaria: Evaluación de diferentes algoritmos de Machine Learning para su predicción," *Revista de Ciencias Sociales*, vol. 28, no. 3, pp. 362–375, Jul. 2022. <https://doi.org/10.31876/rsc.v28i3.38480>



- [24] O. D. Castrillón-Gómez, W. Sarache, S. Ruiz-Herrera, "Predicción de las principales variables que conllevan al abandono estudiantil por medio de técnicas de minería de datos," *Formación universitaria*, vol. 13, no. 6, pp. 217-228, 2020. <https://doi.org/10.4067/S0718-50062020000600217>
- [25] R. Mosquera, O. D. Castrillón, L. Parra, "Predicción de Riesgos Psicosociales en Docentes de Colegios Públicos Colombianos utilizando Técnicas de Inteligencia Artificial," *Información Tecnológica*, vol. 29, no. 4, pp. 267-280, 2018. <https://doi.org/10.4067/S0718-07642018000400267>
- [26] S. Kino, Y.-T. Hsu, K. Shiba, Y.-S. Chien, C. Mita, I. Kawachi, A. Daoud, "A scoping review on the use of machine learning in research on social determinants of health: Trends and research prospects," *SSM Popul Health*, vol. 15, e100836, 2021. <https://doi.org/10.1016/J.SSMPH.2021.100836>
- [27] L. D. Polero, C. M. Garmendia, R. E. Echegoyen, A. Alves de Lima, F. Bertón, F. Lambardi, P. Ariznavarreta, R. Campos, J. P. Costabel, "A Machine Learning Algorithm for Risk Prediction of Acute Coronary Syndrome (ANGINA)," *Revista Argentina de Cardiología*, vol. 88, pp. 9-13, 2020. <https://doi.org/10.7775/rac.v88.i1.17193>
- [28] I. L. Acosta-Guzman, E. A. Varela-Tapia, C. I. Acosta-Varela, J. D. Tumbaco-Bravo, "Body mass index (BMI) prediction using support vector machine algorithms and AI machine learning decision trees," in *Vigesima Primera Conferencia Iberoamericana en Sistemas, Cibernética e Informática*, 2022, pp. 18-23. <https://doi.org/10.54808/CISCI2022.01.18>
- [29] A. G. Quintanilla, N. M. Medina, J. Sulla-Torres, "Prediction of breast cancer through biomarkers using machine learning," in *18th LACCEI International Multi-Conference for Engineering, Education and Technology*, 2020. <https://doi.org/10.18687/LACCEI2020.1.1.514>
- [30] J. Tan, C. Ma, C. Zhu, Y. Wang, X. Zou, H. Li, J. Li, Y. He, C. Wu, "Prediction models for depression risk among older adults: systematic review and critical appraisal," *Ageing Research Reviews*, vol. 83, e101803, 2023. <https://doi.org/10.1016/j.arr.2022.101803>
- [31] A. A. Lara, "Prediction of psychosocial occupational risk in urban transport applying machine learning techniques," *Revista Ibérica de Sistemas e Tecnologias de Informação*, no. E37, pp. 153-165, 2020.
- [32] W. Sun, et al., "Towards artificial intelligence-based learning health system for population-level mortality prediction using electrocardiograms," *NPJ Digital Medicine*, vol. 6, no. 1, e21, 2023. <https://doi.org/10.1038/s41746-023-00765-3>
- [33] M. Molina-Calderon et al., "Application of Machine Learning for the Prediction of Covid19 Through Classification Techniques and Supervised Learning," in *20th LACCEI International Multi-Conference for Engineering, Education, and Technology*, 2022. <https://doi.org/10.18687/LACCEI2022.1.1.425>
- [34] S. A. Sadegh-Zadeh, M. Bahrami, A. Najafi, M. Asgari-Ahi, R. Champion, A. M. Hajiyavand, "Evaluation of COVID-19 pandemic on components of social and mental health using machine learning, analysing United States data in 2020," *Frontiers in Psychiatry*, vol. 13, e933439, 2022. <https://doi.org/10.3389/fpsy.2022.933439>
- [35] K. Todorovic, E. O'Leary, K. P. Ward, P. P. Devarasetty, S. J. Lee, M. Knox, E. Andari, "Prevalence, increase and predictors of family violence during the COVID-19 pandemic, using modern machine learning approaches," *Frontiers in Psychiatry*, vol. 13, e883294, 2022. <https://doi.org/10.3389/fpsy.2022.883294>
- [36] M. Hannes, C. Rauh, "Using past violence and current news to predict changes in violence," *International Interactions*, vol. 48, no. 4, pp. 579-596, 2022. <https://doi.org/10.1080/03050629.2022.2063853>
- [37] D. E. Goin, K. E. Rudolph, J. Ahern, "Predictors of firearm violence in urban communities: A machine-learning approach," *Health and Place*, vol. 51, pp. 61-67, 2018. <https://doi.org/10.1016/j.healthplace.2018.02.013>
- [38] J. Yin, "Crime Prediction Methods Based on Machine Learning: A Survey," *Computers, Materials and Continua*, vol. 74, no. 2, pp. 4601-4629, 2023. <https://doi.org/10.32604/cmc.2023.034190>
- [39] L. Bennett Moses, J. Chan, "Algorithmic prediction in policing: assumptions, evaluation, and accountability," *Policing and Society*, vol. 28, no. 7, pp. 806-822, 2018. <https://doi.org/10.1080/10439463.2016.1253695>
- [40] A. Cáceres, "Rol de la tecnología en la prevención de la violencia de género. El caso de Qispy," Master Thesis, Pontificia Universidad Católica del Perú, Perú, 2021. <http://hdl.handle.net/20.500.12404/20848>
- [41] M. F. Canabal Benito, R. Toro Flores, S. López Ongil, C. López Ongil, "Neurotransmisores para mejorar la detección de situaciones de peligro en víctimas de violencia de género," in *V Congreso Internacional de Jóvenes Investigadorxs con perspectiva de género*, 2020, pp. 314-322. <http://hdl.handle.net/10016/32409>
- [42] L. F. Macedo Quiñones, G. T. Chávez Tarazona, "Aplicación de redes neuronales artificiales sobre la violencia de la mujer por su pareja según la encuesta demográfica y de salud familiar, Endes 2016," Grade

## Machine Learning Applied to Gender Violence: A Systematic Mapping Study

- Thesis, Universidad Nacional Santiago Antúnez de Mayolo, Perú, 2018. <http://repositorio.unasam.edu.pe/handle/UNASAM/2885>
- [43] A. Farouk, A. Nasser, "El árbol de decisión en el análisis de la convivencia en casos de víctimas mortales de violencia de género en España," *Journal of Science and Research*, vol. 5, no. 4, pp. 109–119, 2020. <https://doi.org/10.5281/zenodo.4161179>
- [44] I. Rodríguez-Rodríguez, J. V. Rodríguez, D. J. Pardo-Quiles, P. Heras-González, I. Chatzigiannakis, "Modeling and Forecasting Gender-Based Violence through Machine Learning Techniques," *Applied Sciences*, vol. 10, no. 22, e8244, 2020. <https://doi.org/10.3390/APP10228244>
- [45] C. M. Homan, J. Nicolas Schrading, R. W. Ptucha, C. Cerulli, C. O. Alm, "Quantitative Methods for Analyzing Intimate Partner Violence in Microblogs: Observational Study," *Journal of Medical Internet Research*, vol. 22, no. 11, p. e15347, 2020. <https://doi.org/10.2196/15347>
- [46] J. M. Kafka, M. D. Fliss, P. J. Trangenstein, L. McNaughton Reyes, B. W. Pence, K. E. Moracco, "Detecting intimate partner violence circumstance for suicide: development and validation of a tool using natural language processing and supervised machine learning in the National Violent Death Reporting System," *Injury Prevention*, vol. 29, no. 2, pp. 134-141, Dec. 2022. <https://doi.org/10.1136/IP-2022-044662>
- [47] L. B. Amusa, A. V. Bengesai, H. T. A. Khan, "Predicting the Vulnerability of Women to Intimate Partner Violence in South Africa: Evidence from Tree-based Machine Learning Techniques," *Journal of Interpersonal Violence*, vol. 37, no. 7-8, pp. 5228-5245, 2020. <https://doi.org/10.1177/0886260520960110>
- [48] Á. González-Prieto, A. Brú, J. C. Nuño, J. L. González-Álvarez, "Hybrid machine learning methods for risk assessment in gender-based crime," *Knowledge-Based Systems*, vol. 260, e110130, 2023. <https://doi.org/10.1016/j.knosys.2022.110130>
- [49] L. McDougal, N. Dehingia, N. Bhan, A. Singh, J. McAuley, A. Raj, "using machine learning to explore factors associated with marital sexual violence in a cross-sectional study from India," *BMJ Open*, vol. 11, no. 12, e053603, Dec. 2021. <https://doi.org/10.1136/bmjopen-2021-053603>
- [50] A. Raj, N. Dehingia, A. Singh, J. McAuley, L. McDougal, "Machine learning analysis of non-marital sexual violence in India," *EClinicalMedicine*, vol. 39, e101046, 2021. <http://doi.org/10.1016/j.eclinm.2021.101046>
- [51] N. Dehingia, A. K. Dey, L. McDougal, J. McAuley, A. Singh, A. Raj, "Help seeking behavior by women experiencing intimate partner violence in india: A machine learning approach to identifying risk factors," *Plos One*, vol. 17, no. 2, e0262538, 2022. <https://doi.org/10.1371/journal.pone.0262538>
- [52] S. Lopez-Larrosa et al., "Using Machine Learning Techniques to Predict Adolescents' Involvement in Family Conflict," *Social Science Computer Review*, vol. 0, no. 0, e01, 2022. <https://doi.org/10.1177/08944393221084064>
- [53] J. Xue, J. Chen, C. Chen, R. Hu, T. Zhu, "The hidden pandemic of family violence during COVID-19: Unsupervised learning of tweets," *Journal of Medical Internet Research*, vol. 22, no. 11, e24361, 2020. <https://doi.org/10.2196/24361>
- [54] W. E. Leal, A. R. Piquero, J. Kurland, N. L. Piquero, E. L. Gloyd, "A Case Study of Family Violence During COVID-19 in San Antonio," *Crime Delinq*, vol. 68, no. 8, pp. 1161-1182, Jul. 2022. <https://doi.org/10.1177/00111287211064781>
- [55] M. M. Hossain et al., "Prediction on Domestic Violence in Bangladesh during the COVID-19 Outbreak Using Machine Learning Methods," *Applied System Innovation*, vol. 4, no. 4, e77, Oct. 2021. <https://doi.org/10.3390/asi4040077>
- [56] M. Clur, "Violencia doméstica en Argentina: un modelo de evaluación de riesgos aplicando técnicas de machine learning." Grade Thesis, Instituto Tecnológico de Buenos Aires, Argentina, 2022. <http://ri.itba.edu.ar/handle/123456789/3851>