TELOS: Revista de Estudios Interdisciplinarios en Ciencias Sociales
ISSN 1317-0570 / ISSN (e) 2343-5763
Vol. 28(3). Septiembre-Diciembre 2026

REVISIÓN SISTEMÁTICA DE MODELOS DE APRENDIZAJE AUTOMÁTICO PARA LA ESTIMACIÓN DEL TIEMPO DE DESAPARICIÓN DE LAS PERSONAS

Johanna Enith Aguilar-Reyes*
https://orcid.org/0000-0002-1230-2503
Tania Paulina Morocho-Barrionuevo**
https://orcid.org/0000-0002-1019-6049
Liliana Alejandra Funes-Samaniego***
https://orcid.org/0000-0002-5364-0699
Jorge Juan Zavaleta-Gavidia****
https://orcid.org/0000-0002-4747-8613

RECIBIDO: 27/01/2026 / **ACEPTADO**: 13/08/2026 / **PUBLICADO**: 16/09/2026

Cómo citar: Aguilar-Reyes, J., Morocho-Barrionuevo, T., Funes-Samaniego, L., Zavaleta-Gavidia, J. (2026). Revisión sistemática de modelos de aprendizaje automático para la estimación del tiempo de desaparición de las personas. *Telos: Revista de Estudios Interdisciplinarios en Ciencias Sociales, 28*(3), 716-726. https://doi.org/10.36390/telos283.26

RESUMEN

Este estudio presenta una revisión sistemática sobre la aplicación de modelos de aprendizaje automático al análisis de casos de personas desaparecidas, enfocándose en la predicción temporal, la caracterización del fenómeno social y comparación del desempeño algorítmico. Siguiendo la metodología PRISMA, se identificaron inicialmente 72 estudios, de los cuales solo 12 cumplieron los criterios de inclusión relacionados con pertinencia, calidad metodológica y disponibilidad de información técnica. Los artículos publicados entre 2020 y 2025 evidencian un aumento sostenido en el uso de Machine Learning para abordar problemáticas sociales complejas, con mayor concentración en América y Europa, especialmente en Colombia, Ecuador, España y Polonia. Los modelos analizados comprenden métodos supervisados (Random Forest, Árboles de Decisión, KNN, SVM), redes neuronales profundas (CNN, MLP, LSTM, GRU) y modelos econométricos (VAR, VECM), aplicados según la naturaleza de los datos. A partir del análisis integral de los estudios, se recomienda el uso de modelos basados en árboles (Random Forest y Árboles de Decisión) cuando se busca estabilidad, interpretabilidad relativa y un rendimiento equilibrado; redes neuronales profundas cuando los datos son no lineales, secuenciales o visuales y SVM cuando la precisión es prioritaria y la estructura del conjunto de datos es clara y bien definida. Esta revisión evidencia que el aprendizaje automático representa una herramienta prometedora para fortalecer los procesos de búsqueda e investigación de personas desaparecidas, aunque persisten desafíos asociados a la calidad de los datos, la estandarización de métricas y la necesidad de modelos más explicables para su adopción institucional.

Palabras clave: multivariante, aprendizaje automático, desaparición, predicción, comparación.

Systematic Review of Machine Learning Models for Estimating the Elapsed Time Since Disappearance

ABSTRACT

This study presents a systematic review of the application of machine learning models to the analysis of missing persons cases, focusing on time prediction, characterization of the social phenomenon, and comparison of algorithmic performance. Following the PRISMA methodology, 72 studies were initially identified, of which only 12 met the inclusion criteria related to relevance, methodological quality, and availability of technical information. The articles, published between 2020 and 2025, reveal a sustained increase in the use of machine learning to address complex social problems, with a greater concentration in the Americas and Europe, especially in Colombia, Ecuador, Spain, and Poland. The models analyzed include supervised methods (Random Forest, Decision Trees, KNN, SVM), deep neural networks (CNN, MLP, LSTM, GRU), and econometric models (VAR, VECM), applied according to the nature of the data. Based on a comprehensive analysis of the studies, tree-based models (Random Forest and Decision Trees) are recommended when stability, relative interpretability, and balanced performance are desired; deep neural networks when the data are nonlinear, sequential, or visual; and SVMs when accuracy is paramount and the dataset structure is clear and well-defined. This review shows that machine learning is a promising tool for strengthening the search for and investigation of missing persons, although challenges remain regarding data quality, the standardization of metrics, and the need for more explainable models for institutional adoption.

Keywords: multivariate, machine learning, disappearance, prediction, comparison.

Introducción

El aprendizaje automático (Machine Learning, ML) constituye una de las áreas más dinámicas de la inteligencia artificial, integrando fundamentos de la estadística, la informática y la lógica matemática (Mitchell, 1997; Russell y Norvig, 2020; Forero y Negre, 2024; Costa et al., 2025). Desde sus orígenes en la década de 1950 y su consolidación conceptual en los años ochenta, con aportes fundacionales como los de Michalski et al. (1983), el ML ha experimentado un crecimiento sostenido, transformándose en un eje

* Ingeniera en Estadística Informática por la Escuela Superior Politécnica de Chimborazo – Ecuador. Máster Universitario en Estadística Aplicada por la Univerisdad Politécnica Estatal de Carchi – Ecuador. Curso de Doctorado en la Universidad Nacional de Tumbes, Tumbes – Perú. Docente en la Escuela Superior Politécnica de Chimborazo – Ecuador. johannae.aguilar@espoch.edu.ec
** Biofísica por la Escuela Superior Politécnica de Chimborazo – Ecuador. Máster Universitario en Estadística Aplicada por la Universidad de Granada – España. Docente en la Escuela Superior Politécnica de Chimborazo – Ecuador. tpaulina.morochob@espoch.edu.ec
*** Doctor en Ciencias Económicas por la Universidad de La Habana – Cuba. Docente en la Escuela Superior Politécnica de Chimborazo – Ecuador alejandra.funes@espoch.edu.ec
**** Doctor en Ingeniería de Sistemas y Computación (PESC-UFRJ) y especialización en Inteligencia Artificial. Docente investigador en la Universidad Federal Rural do Rio de Janeiro – Brasil. jorge.zavaleta@ppgi.ufrj.br

CC BY

fundamental de la ciencia de datos por su capacidad para modelar grandes volúmenes de información y patrones complejos (Jordan y Mitchell, 2015; Lemus y Pérez, 2020).

En términos generales, el ML busca desarrollar sistemas capaces de aprender a partir de los datos mediante la identificación automatizada de relaciones y estructuras subyacentes. Sus métodos se agrupan principalmente en aprendizaje supervisado, no supervisado y por refuerzo, permitiendo abordar problemas de clasificación, regresión, agrupamiento, reducción de dimensionalidad y toma de decisiones secuenciales (Goodfellow et al., 2016; Barbierato y Gatti, 2024).

A pesar de su madurez metodológica, el ML enfrenta desafíos importantes, como la limitada interpretabilidad de algunos modelos que operan como verdaderas "cajas negras" y la dificultad de incorporar explícitamente la incertidumbre en las predicciones (Lipton, 2018; Cabanillas et al., 2024; Bagheri et al., 2025). No obstante, su capacidad para representar comportamientos no lineales y su adaptabilidad a distintos tipos de datos lo convierten en una herramienta particularmente valiosa para abordar problemas sociales de alta complejidad.

Uno de estos problemas es la desaparición de personas que es un fenómeno social crítico en diversos países, especialmente en América Latina e implica consecuencias legales, sociales y humanitarias de gran magnitud (Robledo y Querales, 2020; Camacho, 2024;). La gestión oportuna de estos casos requiere herramientas analíticas capaces de estimar parámetros clave, como la duración de la desaparición o la probabilidad de localización, con el fin de optimizar la asignación de recursos y mejorar la efectividad de los procesos de búsqueda. En los últimos años se han desarrollado múltiples estudios que aplican ML al análisis de desapariciones incluyendo modelos de clasificación, predicción temporal y análisis genético forense, pero aún no existe una síntesis sistemática que integre y compare de manera rigurosa los enfoques, arquitecturas, métricas y fuentes de datos empleados.

En este contexto, la presente revisión sistemática, desarrollada bajo las directrices de la metodología PRISMA (Quispe et al., 2021), tiene como objetivo consolidar y analizar críticamente la evidencia disponible sobre los modelos de aprendizaje automático aplicados a casos de personas desaparecidas. La revisión abarca estudios publicados entre 2020 y 2025 en distintas regiones del mundo, evaluando la pertinencia de los métodos utilizados, las métricas de evaluación reportadas, la calidad de las fuentes de datos y el potencial de los modelos para la estimación temporal y la caracterización predictiva de las desapariciones.

Metodología

La revisión sistemática fue realizada siguiendo la metodología PRISMA la cual también ha sido propuesta por otros autores (Barquero, 2022; Chocobar y Barreda, 2025) y sigue un proceso relativamente estandarizado como se observa en el cuadro 1. Allí se ilustra que este procedimiento se enfoca en localizar, seleccionar, evaluar, sintetizar y evidenciar la disponibilidad de modelos multivariantes junto a técnicas de aprendizaje automático aplicadas hacia estimaciones temporales relacionadas en casos de desaparición.

Cuadro 1.
Proceso de revisión sistemática

Fase	Resultado
Identificación	Registros identificados (n = 72)
Aplicación de cribado (filtros lingüísticos, tipo de publicación periodo temporal)	Registros eliminados por duplicación e irrelevancia 50 artículos
Registros evaluados / incluidos	22 artículos
Registros evaluados / excluidos por no cumplir los criterios de inclusión a conveniencia de la investigación	10 artículos
Inclusión final	Publicaciones incluidas en la revisión (n = 12).

De manera más concreta, esta metodología se operacionalizó mediante una serie de fases que, descritas secuencialmente, se desglosan de la siguiente manera:

Formulación de las preguntas de investigación

- (P1): ¿Cuántos estudios se publicaron en los últimos 10 años sobre la estimación de la duración de personas desaparecidas mediante Aprendizaje Automático?
- (P2): ¿Cuáles son los algoritmos o técnicas de Aprendizaje Automático que se emplean para estimar la duración de personas desaparecidas?
- (P3): ¿Qué tipo de variables se ha concluido que son relevantes para la predicción de la duración de personas desaparecidas en base a la revisión sistemática realizada?
- (P4): ¿Cómo se han recolectado los datos para efectuar los algoritmos de Aprendizaje Automático?

Identificación de los estudios

Se efectuó una búsqueda exhaustiva dentro bases académicas reconocidas: Scopus, Web of Science, IEEE Xplore, SpringerLink y Google Scholar utilizando combinaciones léxicas tanto en inglés como en español tales como: "missing persons", "time prediction", "duration estimation", "machine learning", "multivariate models", "desaparición personas" y "modelos predictivos". En esta fase inicial se identificaron 72 registros.

Cribado

De esos cincuenta[i] (50) registros restantes se evaluaron títulos, resúmenes, palabras claves, donde se verificó pertinencia respecto al tema de investigación y se excluyeron 28 artículos que, aunque abordan temas relacionados con ML, carecen de conexión directa en la estimación de duración de casos de desaparición aportando evidencias empíricas aplicables. De los 22 restantes, se excluyeron 10 por no cumplir los criterios de inclusión. Finalmente quedaron 12 artículos relevantes para la revisión sistemática.

Elegibilidad

De este modo, la muestra documental definitiva quedó conformada por doce artículos, cuya inclusión se justifica en el apartado correspondiente a los criterios de selección.

Criterios de inclusión y exclusión

Criterios de inclusión: Publicaciones entre 2020–2025. Estudios en inglés y español. Aplicación de modelos multivariantes, técnicas aprendidas automáticamente orientadas hacia la estimación, duración, predicción de eventos vinculados a desapariciones. Acceso al texto completo y disponibilidad de resultados medibles.

Criterios de exclusión: Documentos sin respaldo empírico, ensayos teóricos carentes de validación. Estudios aplicados a contextos distintos, ejemplos de predicción industrial sin analogía social, artículos con poca claridad respecto a variables, metodología, métricas y evaluación.

Resultados y hallazgos principales

En el cuadro 2, se presentan los estudios más significativos identificados en la revisión, seleccionados por su aporte metodológico, desempeño de modelos y aplicabilidad al contexto de desapariciones.

Cuadro 2.

Estudios seleccionados

N.-	Autor-País	Conjunto de datos	Arquitectura del modelo	Métrica de evaluación
1	Antkiewicz et al., (2020) Polonia.	Casos de personas desaparecidas (Polonia)	Árboles de decisión, Redes neuronales, Naive Bayes	Reportado, valor exacto no especificado
2	Mariño (2025), Colombia.	Series temporales multivariadas económicas	VAR, VECM, MLP, LSTM, GRU	Se reporta ECM, EAM, RECM
3	(Melián 2021), España.	Variables genéricas de clasificación	KNN, SVM, Redes neuronales, Árboles de decisión	SVM mostró mayor precisión ECM
4	(Moreno, 2008), Colombia.	Series de tiempo	Regresión lineal, exponencial, parabólica	Se reporta ECM
5	Aybar et al. (2026), Perú	Casos desapariciones en Perú	Árboles de decisión, Redes neuronales	Reportado, valor exacto no especificado
6	Nirmala et al. (2024), India.	Bases de imágenes públicas (LFW, VGGFace2) y CCTV	CNN (VGG16, ResNet), SVM, KNN	Accuracy, Precision, Recall, F1-score
7	Espinoza y Colina (2025), Ecuador.	68,072 reportes oficiales de personas desaparecidas (DINASED)	XGBoost + SMOTE + SHAP	Accuracy, AUC, Recall, MAE
8	(Delahoz y Mendoza (2021), Colombia.	6,202 reportes de personas desaparecidas	Árboles, Random Forest, Regresión Logística	Accuracy, Sensibilidad, Especificidad
9	Ruiz et al. (2022), Colombia.	Casos de adultos mayores desaparecidos	KNN, Árboles, Random Forest	AUC, Accuracy, F1-score
10	Antkiewicz et al. (2020), Polonia.	Registros administrativos de desaparecidos	Random Forest, SVM, Naive Bayes	Accuracy, Matriz de confusión
11	Huang et al.(2022), Estados Unidos.	Perfiles genéticos con alto error	Clasificación jerárquica, Random Forest, SVM	Precision, Recall, tasa de error
12	Mohamed Ali, et al. (AIP, 2025), India.	Datos de avistamientos, localizaciones y reportes	Redes neuronales + modelos predictivos	Accuracy, Recall, reducción de tiempo

Principales hallazgos

Los estudios realizados en Polonia por Antkiewicz et al. (2020) muestran que los modelos clásicos como árboles de decisión y Naive Bayes ofrecen un rendimiento aceptable en la clasificación de patrones de desaparición, aunque las redes neuronales superan a estos métodos cuando se requiere capturar relaciones no lineales (Carvajal, 2024). El trabajo destaca que la calidad de los registros administrativos tiene un impacto directo en la precisión de los modelos, sugiriendo la necesidad de datos más estandarizados para optimizar el desempeño predictivo.

En Colombia, Delahoz y Mendoza (2021) concluyen que Random Forest supera a modelos como KNN y árboles de decisión en la predicción de características asociadas a casos de desaparición. Estos autores, al igual que Di Martino y Del Mastro (2023) plantean

[i] Estos cincuenta registros corresponden a lo expuesto en el Cuadro 1, específicamente a "Proceso de revisión sistemática", Sección "Aplicación de cribado".

que la heterogeneidad de variables sociodemográficas y contextuales fortalece el desempeño de los modelos basados en árboles, especialmente cuando el conjunto de datos presenta niveles de ruido o falta de completitud.

Por su parte, Mariño (2024) compara diversos modelos de series temporales multivariadas aplicados a datos económicos, encontrando que los modelos basados en redes neuronales recurrentes (LSTM y GRU) presentan un menor error cuadrático medio frente a técnicas econométricas como VAR y VECM. El estudio evidencia que los métodos modernos de deep learning son más eficientes para capturar dependencias dinámicas complejas y no lineales en datos secuenciales.

En España, Melián (2021) demuestra que SVM obtiene el mejor desempeño entre varios algoritmos de clasificación. El estudio concluye que, frente a variables genéricas de clasificación, SVM maximiza el margen de separación y ofrece mayor precisión, superando a modelos como KNN, redes neuronales y árboles de decisión, especialmente cuando el problema presenta fronteras no lineales definidas. En esa apreciación, también se alinean autores como Poveda et al (2024), para quienes el SVM les saca ventaja a otros instrumentos clasificatorios similares.

Moreno (2008) analiza series de tiempo univariadas en Colombia y concluye que los modelos de regresión exponencial ofrecen mejor ajuste frente a los lineales y parabólicos, lo cual es consistente con patrones de crecimiento acelerado presentes en los datos. Este resultado sugiere que la elección del modelo debe considerar la estructura temporal y la dirección de las tendencias observadas: aspecto que coincide con estudios como el de Bastidas et al. (2020), en el que se destaca la importancia del análisis de los métodos de correlación no lineal.

En el ámbito de reconocimiento automático, Nirmala et al. (2024) demuestran que los modelos basados en redes neuronales convolucionales como VGG16 y ResNet, combinados con SVM y KNN, alcanzan altos niveles de precisión, recall y F1-score en la identificación de personas desaparecidas El estudio evidencia que la integración de deep learning en sistemas de vigilancia puede reducir significativamente los tiempos de búsqueda.

En el contexto ecuatoriano, Espinoza y Colina (2025) llevaron a cabo un estudio aplicando modelos de machine learning sobre un conjunto de más de 68 000 reportes oficiales de personas desaparecidas. Los autores concluyen que el algoritmo XGBoost, combinado con un procedimiento de balanceo de clases mediante SMOTE y con técnicas de explicabilidad basadas en SHAP, proporciona el mejor rendimiento predictivo, al tiempo que favorece una interpretación más robusta y transparente de las características relevantes en el modelo.

Los autores resaltan la identificación de variables críticas en la resolución de casos, aunque advierten que los modelos de regresión para estimar tiempo de desaparición aún carecen de precisión por la falta de variables contextuales. Complementando este enfoque, Delahoz y Mendoza (2021) muestran nuevamente que Random Forest y la regresión logística logran altas tasas de sensibilidad y exactitud en la predicción del estado de personas desaparecidas. El estudio subraya la importancia de incorporar variables sociales y psicológicas para mejorar los resultados predictores.

El trabajo de Ruiz et al. (2022) enfocado en adultos mayores desaparecidos resalta que los algoritmos de KNN, árboles de decisión y Random Forest alcanzan altos valores de AUC y F1-score, siendo útiles para evaluar la probabilidad de éxito en la búsqueda. Variables como edad, situación médica y tiempo desde la desaparición se identifican como factores determinantes para las predicciones.

En Polonia, Antkiewicz et al. (2020) demuestran que Random Forest y SVM superan a Naive Bayes en la predicción de razones de desaparición. El estudio, basado en registros administrativos, evidencia que los métodos robustos basados en árboles y márgenes obtienen mejor rendimiento cuando los datos son estructurados y abundantes.

En Estados Unidos, Huang et al. (2022) desarrollan un modelo jerárquico de clasificación para casos de identificación genética afectados por errores de genotipado. Los resultados muestran que la combinación de múltiples medidas genéticas y algoritmos como Random Forest y SVM mantiene altos niveles de precisión incluso en escenarios de ruido alto, lo que los convierte en herramientas valiosas en contextos forenses de personas desaparecidas.

Finalmente, el estudio AI-assisted Search (Mohamed Ali et al., 2025) muestra que el uso de redes neuronales y modelos predictivos permite integrar datos de avistamientos, localizaciones y patrones de movilidad para reducir significativamente los tiempos de búsqueda. Los autores destacan que esta integración multimodal representa una vía prometedora para mejorar la eficiencia operativa de equipos de búsqueda y rescate.

Respuestas para las preguntas de investigación formuladas

Respuesta a la P1, Estudios publicados: se identificaron 12 estudios publicados entre 2020 y 2025 provenientes principalmente de Polonia, Colombia, India, Ecuador, España y Estados Unidos**,** lo que evidencia un interés global y multidisciplinario en el uso de técnicas de aprendizaje automático para abordar problemáticas asociadas a personas desaparecidas.

Respuesta a la P2, Modelos utilizados: Los 12 artículos analizados evidencian un amplio espectro de modelos de aprendizaje automático aplicados al estudio de desapariciones, reconocimiento, predicción y análisis de patrones asociados. Sobresalen los árboles de decisión y los métodos ensemble, especialmente Random Forest, empleados en estudios de Polonia, Colombia y Ecuador para clasificar causas de desaparición y predecir estados o probabilidades de localización (Antkiewicz et al., 2020; Delahoz y Mendoza 2021; Aybar et al., 2026).

Asimismo, se identifican redes neuronales artificiales y profundas, que incluyen modelos como MLP, así como arquitecturas avanzadas LSTM y GRU, utilizados principalmente en contextos de series temporales y predicción de dinámicas complejas (Camejo et al., 2020; Mariño, 2024; Mohamed Ali et al., 2025). En paralelo, varios estudios implementan clasificadores clásicos como SVM, KNN y Naive Bayes**,** los cuales muestran desempeños competitivos en tareas específicas de reconocimiento facial y análisis de registros estructurados (Melián, 2021; Nirmala et al., 2024; Ruiz et al., 2022).

También se incorporan modelos estadísticos y econométricos, entre ellos VAR, VECM y regresiones lineales, exponenciales y parabólicas, aplicados principalmente en investigaciones basadas en series de tiempo (Moreno, 2008). Finalmente, se resaltan enfoques híbridos que integran técnicas estadísticas con algoritmos de ML de última generación, como XGBoost combinado con SMOTE y SHAP para el análisis de grandes bases de datos de desaparecidos en Ecuador (Espinoza y Colina, 2025) o modelos jerárquicos para datos genéticos con alto error (Huang et al., 2022) En conjunto, los hallazgos indican que los modelos ensemble

especialmente Random Forest y XGBoost y las redes neuronales profundas (LSTM/GRU y CNN) se destacan por ofrecer los mejores niveles de precisión, estabilidad y capacidad predictiva en los diversos escenarios evaluados.

Respuesta a la P3, Variables consideradas: sobre este aspecto, por la amplitud y diversidad de variables utilizadas en los estudios y, para mejor visualización de ese hallazgo, se construyó una descripción que se muestra en el cuadro 3.

Cuadro 3.

Las variables más usadas.

Variables	Sinopsis de los estudios/autores que las usan
Variables sociodemográficas	Los estudios revisados coinciden en que las variables sociodemográficas constituyen uno de los pilares fundamentales en los modelos predictivos. Entre las más empleadas destacan edad, género, estado civil, nacionalidad y condición médica. Artículos como los de Ruiz et al. (2022) y, Delahoz y Mendoza (2021) muestran que estas variables poseen un fuerte poder discriminante, especialmente en la predicción de estados (encontrado vivo, fallecido, aún desaparecido) o en la identificación de factores de riesgo. La edad, por ejemplo, se ha considerado crítica en casos de personas mayores desaparecidas, mientras que el género aparece asociado a diferentes patrones de desaparición y vulnerabilidad.
Variables espaciales	Aparecen principalmente en estudios realizados en Polonia, India y Ecuador Antkiewicz et al., (2020); Mohamed Ali et al., (2025); Espinoza y Colina, (2025). En investigaciones basadas en registros policiales o sistemas de videovigilancia, las características espaciales permiten trazar rutas probables, generar mapas de calor, clasificar zonas de mayor incidencia y fortalecer modelos de predicción geográfica. Estas variables aportan información crucial para algoritmos de redes neuronales y modelos ensemble.
Variables contextuales e históricas	Se observan especialmente en estudios que combinan bases administrativas y datos complementarios, como en Espinoza y Colina (2025), donde variables relacionadas con la situación socioeconómica o con el historial del caso aumentan la capacidad explicativa del modelo. En estudios forenses, como Huang et al. (2022), variables históricas de laboratorio y calidad de las muestras genéticas también son determinantes.
Variables de series temporales	Los artículos que trabajan con predicciones dinámicas o análisis longitudinales como Mariño (2024) y Moreno (2008), emplean variables relacionadas con tendencias, estacionalidad, tiempo transcurrido, evolución temporal del caso, fecha del reporte y patrones temporales multivariados**.** En estudios de desapariciones humanas, las variables temporales se utilizan para estimar el tiempo de resolución**,** analizar patrones anuales o modelar la progresión del riesgo según el tiempo desaparecido. Su uso permite aplicar modelos como LSTM, GRU, VAR y VECM, los cuales capturan dependencias temporales complejas y mejoran la capacidad predictiva cuando los datos presentan comportamiento secuencial.

Respuesta a la P4, Recolección de datos: Los 12 estudios analizados utilizan bases de datos provenientes de diversas fuentes institucionales, académicas y técnicas. Una parte importante de los artículos emplea registros oficiales de personas desaparecidas, especialmente los generados por policías y fiscalías de países como Polonia, Colombia y Ecuador, incluyendo bases robustas como los registros administrativos polacos, los reportes oficiales colombianos y los 68.072 casos documentados por DINASED en Ecuador. Otros trabajos utilizan series temporales históricas y económicas obtenidas de repositorios públicos y bases estadísticas nacionales, como ocurre en los estudios de Colombia centrados en dinámica temporal.

En el ámbito experimental y académico, algunos artículos recurren a repositorios universitarios y conjuntos de datos genéricos de clasificación, empleados principalmente en investigaciones realizadas en España y Colombia. Para los estudios de visión computacional, las bases provienen de repositorios públicos de imágenes, como LFW y VGGFace2, complementados con vídeos de CCTV recopilados para validación de modelos en India. Finalmente, en el campo forense, se utilizaron bases de datos genéticas con errores de laboratorio del Center for Human Identification de Estados Unidos, mientras que los modelos de localización asistida por IA en India se alimentaron de datos de avistamientos, reportes locales y registros geoespaciales diseñados para evaluar el desempeño del sistema predictivo.

Ahora bien, como cierre de este segmento analítico que incluye autores y estudios referenciados, en el cuadro 4, se describen brevemente los estudios más significativos identificados en la revisión, una vez que han sido seleccionados por su aporte metodológico, desempeño de modelos y aplicabilidad a las desapariciones.

Cuadro 4.

Estudios destacados por sus aportes a esta investigación.

Autores/estudios /enlace	Descripción de aporte
Antkiewicz et al. (2020): Prediction of Missing Persons' Disappearance Reason Using ML	Aborda la predicción de las razones de desaparición de personas mediante modelos de machine learning. Se emplearon árboles de decisión, redes neuronales y Naive Bayes, utilizando como variables principales la edad, género, ubicación e historial de los individuos. Para la evaluación se aplicaron métricas de accuracy y F1-score, destacándose que los árboles de decisión alcanzaron el mejor desempeño predictivo frente a los demás modelos.
Delahoz y Mendoza (2021): Modelo Predictivo para Personas Desaparecidas en Colombia	Propone un modelo predictivo para personas desaparecidas en Colombia, empleando técnicas de machine learning como KNN, árboles de decisión y Random Forest. Se consideraron variables sociodemográficas como edad, género, estado civil y ubicación. La evaluación se realizó mediante métricas de accuracy, precision y recall, evidenciando que el Random Forest alcanzó el mejor desempeño en la predicción de los casos de desaparición

Mariño (2025). Comparación de Modelos Estadísticos y ML para Series Temporales Multivariadas (Colombia)	Se centra en la comparación entre modelos estadísticos tradicionales (VAR y VECM) y técnicas de machine learning aplicadas a series temporales multivariadas en el contexto económico colombiano. Los resultados mostraron que las redes neuronales recurrentes, en especial el modelo LSTM, alcanzaron mayor precisión al reducir los errores de predicción frente a los métodos estadísticos
Melián (2021) Revisión de Modelos de Predicción en Machine Learning (España)	Constituye una revisión de diferentes algoritmos de predicción en machine learning, incluyendo KNN, SVM, redes neuronales y árboles de decisión. Entre los hallazgos más relevantes se destaca el desempeño del SVM, que logró altos niveles de precisión en tareas de clasificación.
Moreno (2008): Predicción con Series de Tiempo y Regresión (Colombia)…	Resalta que, en ciertos escenarios, los métodos clásicos de regresión mantienen vigencia y pueden ofrecer predicciones satisfactorias sin necesidad de recurrir a técnicas más sofisticadas.
Nirmala et al. (2024), India: A Survey of Integrating Deep Learning Based Missing Person Detection Model Into CCTV Systems…	Plantea que la integración de IA en sistemas de videovigilancia puede reducir significativamente el tiempo de localización, aunque requiere infraestructura computacional robusta.
Espinoza y Colina (2025), Ecuador: Predictive Analysis of Missing Persons in Ecuador (2014-2024)	Logra identificar variables críticas en la resolución de casos, aunque también revela que los modelos de regresión para estimar el tiempo de desaparición aún son débiles por la ausencia de datos contextuales detallados.
(Gelvez, et al. , 2022) Prediciendo el crimen en ciudades intermedias: un modelo de "machine learning" en Bucaramanga, Colombia	En este texto, los autores presentan los resultados de un modelo de "machine learning" para predecir el delito en Bucaramanga, una ciudad intermedia de Colombia. Se utilizó el procesamiento de señales para grafos y una adaptación al caso del modelo de vectorización de texto TF-IDF
Ruiz et al. (2022), Alemania / Colombia: Predicting the probability of finding missing older adults based on machine learning…	Destaca que la edad, condición médica y tiempo transcurrido desde la desaparición son factores críticos, así mismo, el estudio aporta evidencia de que el uso de ML puede apoyar a instituciones de búsqueda, especialmente en casos de personas vulnerables.
Antkiewicz et al. (2020), Polonia: Prediction of Missing Persons' Disappearance Reason Using Machine Learning Methods…	Los resultados muestran que Random Forest y SVM superan a Naive Bayes en la predicción de causas de desaparición. La alta exactitud alcanzada se atribuye a la estructura bien definida de los datos policiales. Además, el estudio sugiere que los métodos basados en árboles y márgenes son más adecuados para clasificar eventos de desaparición cuando los datos están bien estructurados
Huang et al. (2022), Estados Unidos: A machine learning approach for missing persons cases with high genotyping errors…	Presenta un método de clasificación jerárquica para identificación genética en casos de personas desaparecidas. Random Forest y SVM lograron alta precision y recall incluso con errores genéticos elevados. Random Forest y SVM lograron alta precision y recall incluso con errores genéticos elevados. El estudio demuestra que los algoritmos modernos pueden corregir ruido significativo en datos forenses, aportando herramientas de apoyo para identificación humana en contextos complejos.
Mohamed Ali, et al. (AIP, 2025), India.	Una aproximación al proceso de manejo de los datos de avistamientos, localizaciones y reportes cuando se trata de personas registradas como desaparecidas

Llegados a este punto, conviene ver los resultados del Anova, tal como se muestra en la tabla 1.

Tabla 1.
Análisis de ANOVA por métrica de evaluación

Métrica	Fuente de variación	Suma de cuadrados (sum_sq)	gl (df)	F	p-value (PR > F)
Accuracy	Modelo	1172,342	5	103,8527	1,8827 × 10^-15
Accuracy	Residual	54,1848	24	—	—
Precision	Modelo	1607,587	5	159,8596	1,31092 × 10^-17
Precision	Residual	48,26998	24	—	—
Recall	Modelo	2323,89	5	325,8161	3,11682 × 10^-21
Recall	Residual	34,2361	24	—	—
F1-score	Modelo	1413,306	5	132,0564	1,19611 × 10^-16
F1-score	Residual	51,371	24	—	—
Especificidad	Modelo	115,8762	5	23,0113	1,95307 × 10^-8
Especificidad	Residual	24,17099	24	—	—

Estos resultados evidencian diferencias estadísticamente significativas entre los modelos en todas las métricas evaluadas (Accuracy, Precision, Recall, F1 y Specificit), lo que confirma que no todos presentan el mismo nivel de desempeño. En términos generales, Decision Tree se posiciona como el modelo más consistente al destacar en exactitud, precisión y F1, seguido muy de cerca por Random Forest**,** que ofrece un rendimiento similar. Por su parte, Gradient Boosting sobresale en Recall**,** mostrando mayor capacidad para identificar correctamente los casos positivos, mientras que Naive Bayes **y** SVM presentan desempeños más bajos en la mayoría de los indicadores.

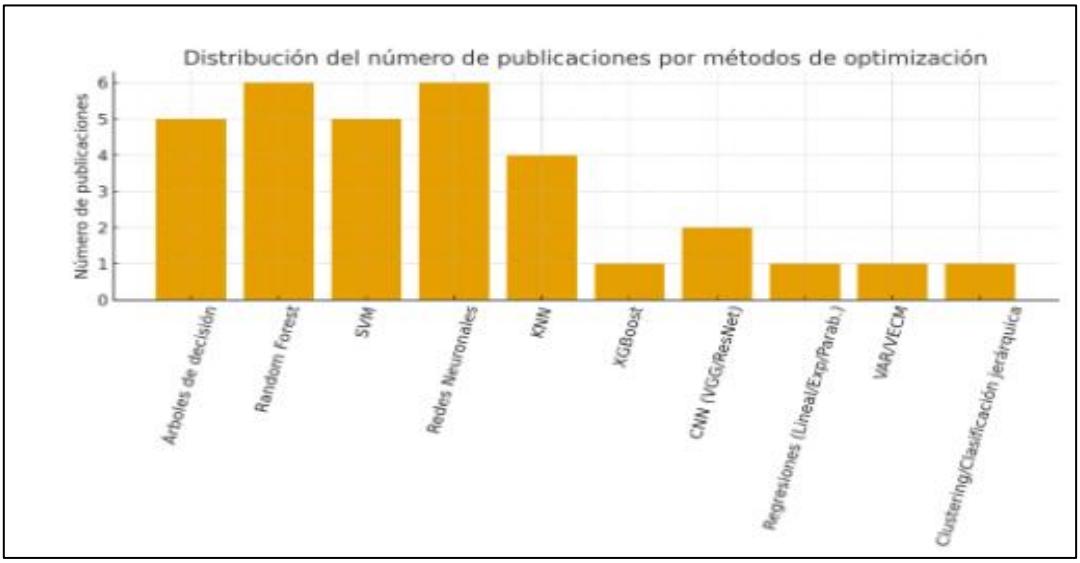


Figura 1. Distribución del número de publicaciones por métodos de optimización

En la Figura 1 se muestra que los métodos más utilizados en las investigaciones revisadas son Random Forest y las Redes Neuronales, lo que evidencia su preferencia en problemas complejos por su alta precisión y capacidad de modelar relaciones no lineales. Les siguen los Árboles de Decisión, SVM y KNN, que –aunque son modelos tradicionales–, continúan empleándose ampliamente por su simplicidad y buen desempeño. En contraste, métodos como XGBoost, regresiones clásicas, VAR/VECM, clustering y CNN específicas aparecen con menor frecuencia, reflejando que su uso depende de contextos más especializados.

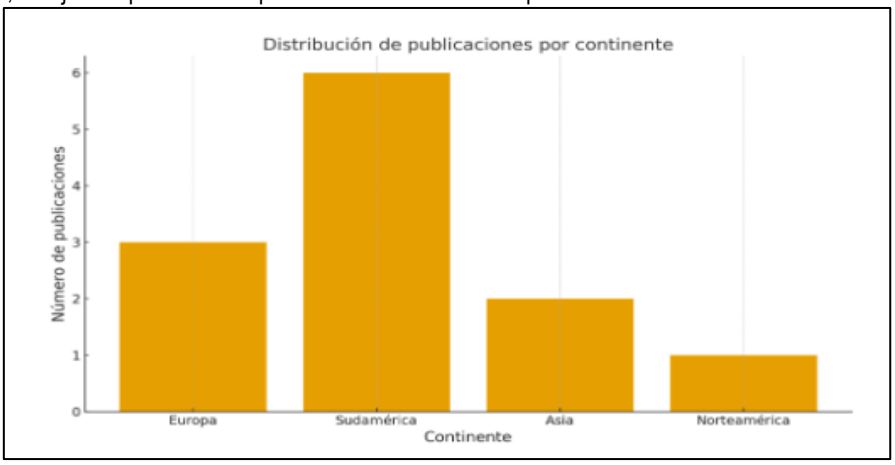


Figura 2. Distribución de artículos por continente.

La figura 2 muestra que Sudamérica concentra la mayor cantidad de publicaciones, con un total de 6 estudios, lo que evidencia un interés significativo de esta región por aplicar modelos de aprendizaje automático al análisis de desapariciones, particularmente en países como Colombia y Ecuador. En segundo lugar se encuentra Europa, con 3 publicaciones, procedentes principalmente de Polonia y España, donde se enfocan en el análisis de registros institucionales y en la experimentación con algoritmos clásicos y avanzados. Asia aparece con 2 publicaciones, destacando investigaciones de India relacionadas con visión computacional y sistemas predictivos basados en inteligencia artificial. Finalmente, Norteamérica registra 1 estudio, proveniente de Estados Unidos, centrado en el uso de datos genéticos forenses.

En conjunto, la distribución evidencia que el mayor volumen de estudios proviene de regiones con problemáticas relevantes en torno a la desaparición de personas como Sudamérica y también de regiones con fuerte desarrollo tecnológico, como Europa, India y Estados Unidos. Esto sugiere que la investigación sobre desapariciones mediante técnicas de machine learning tiene un carácter tanto social como tecnológico, con contribuciones de diversas partes del mundo.

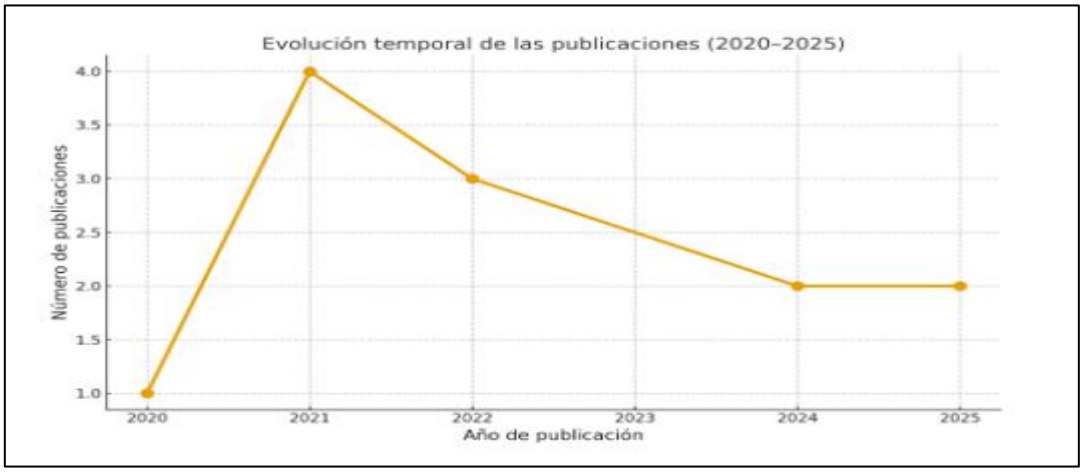


Figura 3. Evaluación temporal de publicaciones 2014-2022

La figura 3 presenta la evolución temporal de las investigaciones sobre aprendizaje automático aplicado al análisis de personas desaparecidas y muestra –además– un crecimiento relevante a partir del año 2021, donde se registra el mayor número de publicaciones de todo el periodo (n = 4). Este incremento coincide con el fortalecimiento global de técnicas de *machine learning* en áreas de seguridad, análisis policial e investigación social. En 2022 se mantiene una producción relativamente alta (n = 3), lo que refleja continuidad en el interés científico por el tema.

Luego, en 2024 y 2025, el número de publicaciones se estabiliza (n = 2 para ambos años), lo que sugiere una consolidación del campo, con trabajos más especializados, como modelos predictivos avanzados, análisis con deep learning y el uso de datos biométricos o genéticos. El año 2020 presenta solo una publicación, lo que indica que la producción científica antes de 2021 era todavía incipiente.

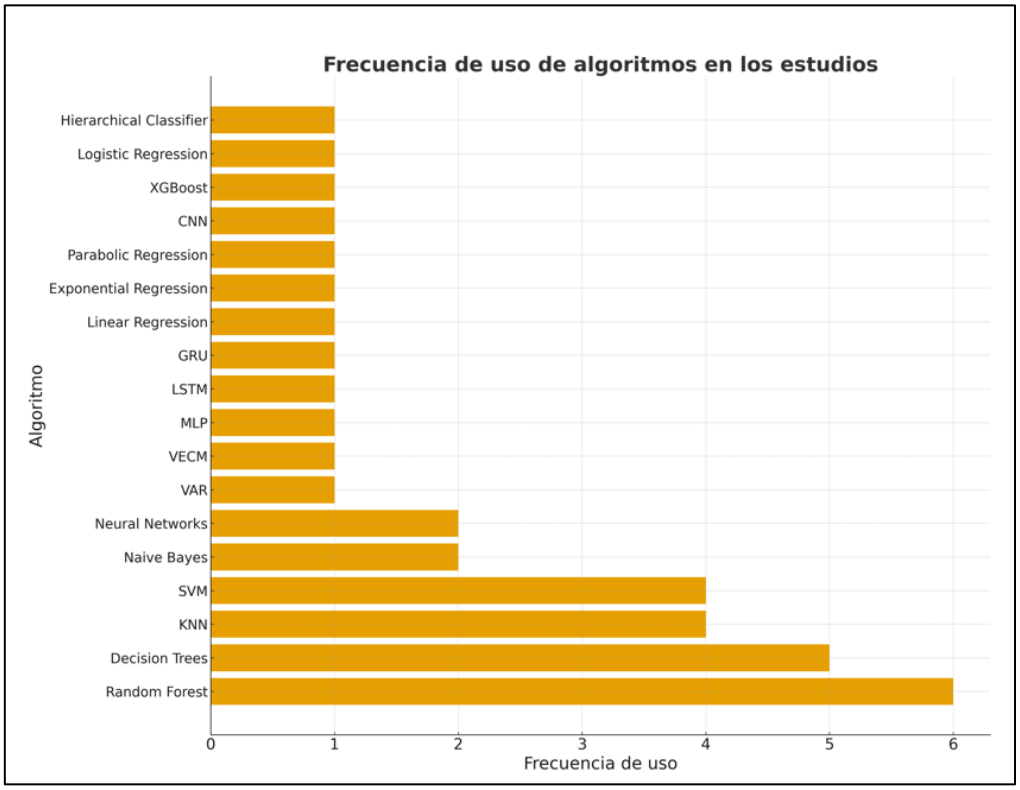


Figura 4. Frecuencia de uso de algoritmo en los estudios

En la figura 4 se observa que, por una parte, Random Forest es el algoritmo más utilizado en los estudios revisados (n = 6), lo que confirma su popularidad por su robustez, capacidad de manejar datos heterogéneos y buen desempeño en problemas de clasificación asociados a personas desaparecidas; por otra parte, le sigue Decision Trees (n = 5), que funciona como modelo base en investigaciones más tradicionales y sirven como referencia para comparar mejoras con algoritmos más avanzados.

Los métodos KNN y SVM aparecen con una frecuencia moderada (n = 4 cada uno), utilizados principalmente en tareas de clasificación de estados o reconocimiento facial, donde su simplicidad y efectividad resultan adecuadas. Algoritmos como Naive Bayes, Neural Networks, Logistic Regression y enfoques específicos (CNN, XGBoost, LSTM, GRU, VAR, VECM) aparecen con menor frecuencia (n = 1 o 2), reflejando su aplicación en estudios más especializados o contextos técnicos particulares, como predicción temporal, análisis genético o visión computacional.

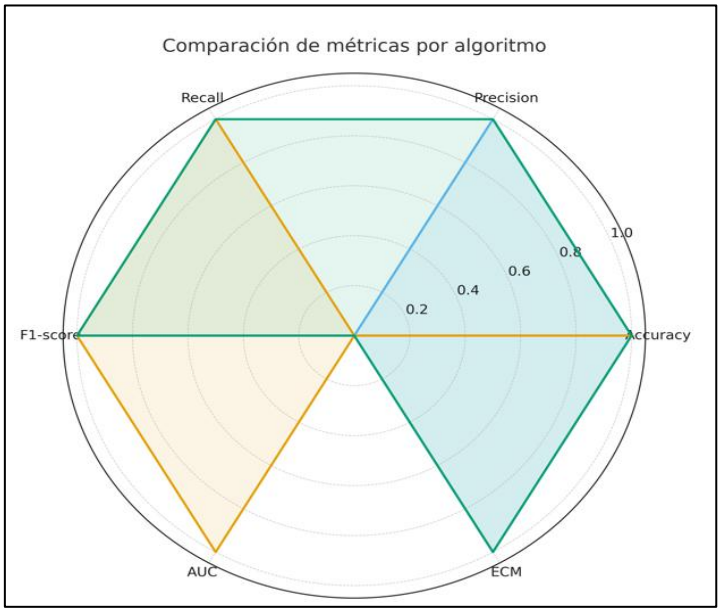


Figura 5. Comparación de métricas por algoritmo

La figura de radar muestra el desempeño de los algoritmos en distintas métricas y, en este caso, revela diferencias importantes en las métricas utilizadas para evaluar cada algoritmo según los 12 estudios analizados. Random Forest muestra la cobertura más amplia de métricas, destacándose por el uso frecuente de *Accuracy*, *Recall*, *F1-score* y *AUC*. Esto evidencia que es el algoritmo más robusto y ampliamente evaluado en problemas de clasificación relacionados con personas desaparecidas, ya que permite medir tanto el rendimiento general como la capacidad de identificar correctamente casos positivos y negativos.

En cuanto al SVM**,** el gráfico muestra un énfasis en *Precision*, *F1-score* y *Accuracy*, así como el uso ocasional del *ECM* en estudios experimentales. Esto indica que SVM se utiliza sobre todo en contextos donde la separación de clases debe ser muy precisa, como en problemas de reconocimiento facial o clasificación con fronteras complejas, privilegiando la minimización de falsos positivos.

Por su parte, las Redes Neuronales (MLP, LSTM, CNN, GRU) abarcan también varias métricas, especialmente *Accuracy*, *Precision*, *Recall* y *F1-score* y –además– emplean métricas de error como el *ECM*, lo cual refleja su aplicación tanto en tareas de clasificación como en predicciones temporales. Su versatilidad se evidencia en la capacidad de emplear métricas continuas y categóricas, mostrando que son los modelos más utilizados para problemas complejos como series de tiempo, visión computacional y análisis de patrones no lineales.

Discusión

Los resultados de esta revisión muestran que el aprendizaje automático se ha convertido en una herramienta central para el análisis de casos de personas desaparecidas, con una adopción creciente desde 2020 y una mayor concentración de estudios entre 2021 y 2025. Esta tendencia coincide con el aumento global en la disponibilidad de datos institucionales y la madurez de modelos predictivos aplicados a fenómenos sociales complejos (Antkiewicz et al., 2020; Delahoz y Mendoza 2021; Tumbaco et al., 2025; Ibarra et al., 2025).

Un hallazgo consistente es que los métodos ensemble, en particular Random Forest, son los algoritmos más utilizados y mejor evaluados en los estudios revisados. Su presencia en seis de los doce artículos y su reporte frecuente con métricas como *Accuracy*, *Recall*, *F1-score* y *AUC* corroboran su robustez frente a datos desbalanceados y heterogéneos, típicos en registros policiales y administrativos de desapariciones (Delahoz y Mendoza, 2021; Antkiewicz et al., 2020; Espinoza y Colina, 2025).

Estos resultados sugieren que Random Forest continúa siendo un modelo altamente confiable para tareas de clasificación en contextos reales. Por otra parte, algoritmos como SVM y KNN se emplean con frecuencia moderada y predominan en estudios donde el objetivo es la clasificación estricta o el reconocimiento de patrones claramente estructurados, como en trabajos de visión computacional con rostros o en bases académicas controladas (Melián, 2021; Nirmala et al., 2024). Las métricas asociadas a estos modelos —principalmente *Precision* y *F1-score*— permiten controlar falsos positivos, un aspecto crítico en tareas de identificación automática.

Las redes neuronales profundas (CNN, MLP, LSTM y GRU) se destacan en investigaciones donde los datos son temporalmente complejos o donde se requiere interpretar información de imágenes. Su presencia en estudios de series temporales y en análisis de CCTV confirma su capacidad para capturar patrones no lineales y dependencias de largo plazo que los modelos estadísticos tradicionales no logran representar adecuadamente (Mariño, 2024; Mohamed Ali et al., 2025). En el campo forense, los modelos neuronales y SVM demostraron precisión incluso bajo condiciones de error genético elevado (Huang et al., 2022), reforzando su relevancia para la identificación humana.

En cuanto a las métricas, se observa una diferenciación clara según el tipo de problema: *Accuracy, Recall, F1-score y AUC* dominan en estudios de clasificación, mientras que *ECM, RECM y EAM* se aplican exclusivamente en predicción temporal (Moreno, 2008). Esta división metodológica sugiere que la selección de métricas está correctamente alineada con los objetivos analíticos de cada estudio, lo cual fortalece la validez de los resultados reportados.

En conjunto, la evidencia revisada indica que los modelos de aprendizaje automático, especialmente Random Forest, XGBoost y las redes neuronales profundas, constituyen enfoques relevantes y prometedores para mejorar la clasificación, predicción e identificación en contextos de desapariciones. Sin embargo, persisten desafíos relacionados con la calidad de los datos, la estandarización de métricas y la interpretabilidad de modelos complejos, aspectos que deben considerarse en futuras investigaciones.

Conclusiones

El uso del aprendizaje automático para el análisis de personas desaparecidas es un campo en crecimiento, con un incremento notable de publicaciones entre 2021 y 2025. Este aumento refleja tanto la disponibilidad de datos institucionales como la necesidad de herramientas predictivas más precisas para apoyar procesos de investigación y búsqueda.

Random Forest emerge como el algoritmo más consistente y robusto, razón por la que es utilizado en la mayoría de los estudios revisados y evaluado con un amplio conjunto de métricas (Accuracy, Recall, F1-score, AUC). Su rendimiento superior confirma que los métodos ensemble son altamente adecuados para manejar datos desbalanceados y heterogéneos, características comunes en los registros de desapariciones.

SVM y KNN continúan siendo alternativas relevantes, especialmente en contextos donde las clases están claramente definidas, como en estudios con datos estructurados o sistemas de reconocimiento facial. Sin embargo, su desempeño suele ser superado por modelos más complejos como Random Forest y XGBoost en bases de datos más amplias o ruidosas. Las redes neuronales profundas (CNN, LSTM, GRU) muestran ventajas significativas en escenarios que involucran imágenes, patrones no lineales o series temporales.

En particular, los modelos LSTM y GRU superan a enfoques econométricos tradicionales (VAR, VECM) en predicción dinámica, mientras que las CNN logran precisión elevada en reconocimiento facial y análisis de CCTV. La selección de métricas es coherente con la naturaleza del problema, utilizándose Accuracy, Recall, F1-score y AUC para clasificación, mientras que ECM, EAM y RECM se aplican en predicción temporal. Esto evidencia rigor metodológico y fortalece la validez de los hallazgos de los estudios revisados.

La diversidad de fuentes de datos incluyendo registros policiales, bases institucionales, repositorios de imágenes, datos genéticos y series temporales demuestra que el fenómeno social de las desapariciones requiere enfoques flexibles y multidimensionales. La calidad y estructura de los datos siguen siendo desafíos críticos que afectan el desempeño de algunos modelos.

Declaración de Conflictos de Interés

No declaran conflictos de interés.

Contribución de autores

Autor	Conceptualización	Curación	Analisis/ Software	Metodología	Recursos	Supervisión/ Validación	Escritura inicial	Escritura final, revisión
1	X	x	x	x	x	x	x	x
2	X		x		x	x		x
3	X				x	x	x	
4	X	x	x	x				x

Financiamiento

Ninguno.

Referencias

Antkiewicz, R., Czuba, C., Gutowski, T., y Pierzchała, D. (2020). Prediction of missing persons' disappearance reason using machine learning methods. In *Proceedings of the 36th International Business Information Management Association Conference (IBIMA)*. https://www.researchgate.net/profile/Tomasz-Gutowski/publication/348835649_Prediction_of_Missing_Persons%27_Disappearance_Reason_Using_Machine_Learning_Methods/links/6012890445851517ef1f3f5e/Prediction-of-Missing-Persons-Disappearance-Reason-Using-Machine-Learning-Methods.pdf?origin=publication_detail&_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uRG93bmxvYWQiLCJwcmV2aW91c1BhZ2UiOiJwdWJsaWNhdGlvbiJ9fQ

Aybar, A., Maehara, R., Benites, L., y Ernesto, M. (2026). Uncovering Disappearance Dynamics in Peru: A Two-Stage Multimodal Framework for Behavioral Segmentation and Voluntariness Prediction in Missing-Person Reports., En *Focus on Artificial Intelligence in Intelligent Systems Design. (4) 1.* https://doi.org/10.1007/978-3-032-20752-4_1

Bagheri, M., Bagheritabar, M., Alizadeh, S., Parizi, M., Matoufinia, P., y Luo, Y. (2025). Machine-learning-powered information systems: A systematic literature review for developing multi-objective healthcare management. *Applied Sciences, 15*(1), Article 296. https://doi.org/10.3390/app15010296

Barbierato, E., y Gatti, A. (2024). The challenges of machine learning: A critical review. *Electronics, 13*(2), Article 416. https://doi.org/10.3390/electronics13020416

Barquero, W. (2022). Análisis PRISMA como metodología para revisión sistemática: Una aproximación general. *Salud en Redes, 8*(1, Suppl.), 339–360. https://doi.org/10.18310/2446-4813.2022v8nsup1p339-360

Bastidas, J., Peña, A., Obando, L., y Montenegro, A. (2020). Importancia de los modelos de regresión no lineales en la interpretación de datos de la COVID-19 en Colombia. *Revista Habanera de Ciencias Médicas, 19*, Article e3309. http://www.revhabanera.sld.cu/index.php/rhab/article/view/3309

Cabanillas, E., Irigoin, U., Lázaro, J., Agurto, C., Esquivel, O., y Alvez, C. (2024). *Ciencia de datos e inteligencia artificial*. Mar Caribe. https://editorialmarcaribe.es/wp-content/uploads/2025/02/SPA-Ciencia-de-datos-e-inteligencia-artificial.pdf

Camacho, J. (2024). Ni vivas ni muertas: Personas desaparecidas en América Latina. Impactos psicosociales en sus familias. *Rupturas, 14*(1), 107–122. https://www.scielo.sa.cr/pdf/rup/v14n1/2215-2989-rup-14-01-107.pdf

Camejo, J., Gonzalez, H., y Morell, C. (2020). Un estudio empírico del modelo de red neuronal MLP para problemas de predicción con salidas múltiples. *Serie Científica de la Universidad de las Ciencias Informáticas, 13*(6), 1–14. https://publicaciones.uci.cu/index.php/serie/article/view/578

Carvajal, C. (2024). Uso de técnicas como la regresión y redes neuronales para anticipar el rendimiento del maíz. *RECIMUNDO, 8*(4), 126–135. https://recimundo.com/index.php/es/article/view/2471

Chocobar, E., y Barreda, R. (2025). Estructuras metodológicas PICO y PRISMA 2020 en la elaboración de artículos de revisión sistemática: Lo que todo investigador debe conocer y dominar. *Ciencia Latina Revista Científica Multidisciplinar, 9*(1), 8525–8543. https://doi.org/10.37811/cl_rcm.v9i1.16491

Costa, J., de Lima, P., Queiroz, M., Loss, L., Lopes, J., de Aguiar, K., y Alves dos Reis, R. (2025). El rol de la inteligencia artificial y el aprendizaje automático en la enseñanza moderna. *Revista Caribeña de Ciencias Sociales, 14*(8), 1–25. https://doi.org/10.55905/rcssv14n8-010

Delahoz, E., y Mendoza, S. (2021). A predictive model for the missing people problem. *Romanian Journal of Legal Medicine, 29*(1), 74–80. https://www.safetylit.org/citations/index.php?fuseaction=citations.viewdetails&citationIds%5B%5D=citjournalarticle_690625_12

Di Martino, F., y Del Mastro, F. (2023). Explainable AI for clinical and remote health applications: A survey on tabular and time series data. *Artificial Intelligence Review, 56*, 5261–5315. https://doi.org/10.1007/s10462-022-10304-3

Espinoza, M., y Colina, A. (2025). Predictive analysis of missing persons in Ecuador (2014–2024). *CEUR Workshop Proceedings, 4055*, 197–210. https://ceur-ws.org/Vol-4055/icaiw_wdea_4.pdf

Forero, W., y Negre, F. (2024). Técnicas y aplicaciones del machine learning e inteligencia artificial en educación: Una revisión sistemática. *RIED-Revista Iberoamericana de Educación a Distancia, 27*(1), 209–253. https://doi.org/10.5944/ried.27.1.37491

Gelvez, Nieto, y Rocha (2022) Prediciendo el crimen en ciudades intermedias: un modelo de "machine learning" en Bucaramanga, Colombia. URVIO, Revista Latinoamericana de Estudios de Seguridad, (34), 83-98. https://doi.org/10.17141/urvio.34.2022.5395

Goodfellow, I., Bengio, Y., y Courville, A. (2016). *Deep learning*. MIT Press.

Huang, M., Liu, M., Li, H., King, J., Smuts, A., Budowle, B., y Ge, J. (2022). A machine learning approach for missing persons cases with high genotyping errors. *Frontiers in Genetics, 13*, Article 971242. https://doi.org/10.3389/fgene.2022.971242

Ibarra, M., Mendoza, E., y Romero, W. (2025). Estudio de modelos de predicción en Big Data. *Ciencia y Educación, 6*(6.1), 337–347. https://www.cienciayeducacion.com/index.php/journal/article/view/zenodo.17051783/1621

Jordan, M. I., y Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*(6245), 255–260. https://doi.org/10.1126/science.aaa8415

Lemus, D., y Pérez, R. (2020). Ciencia de datos y estudios globales: Aportaciones y desafíos metodológicos. *Colombia Internacional, 102*, 41–62. https://doi.org/10.7440/colombiaint102.2020.03

Lipton, Z. C. (2018). The mythos of model interpretability. *Communications of the ACM, 61*(10), 36–43. https://doi.org/10.1145/3233231

Mariño, D. (2025). *Comparación de modelos estadísticos y de machine learning para la estimación de la duración de desapariciones* [Tesis de maestría, Universidad Nacional de Colombia]. https://repositorio.unal.edu.co/items/618ecf2a-a482-48cf-854f-b7cb4ac71a5b

Melián, C. (2021). *Revisión de modelos de predicción aplicados en ciencias sociales mediante machine learning* [Trabajo de grado, Universidad de La Laguna]. http://riull.ull.es/xmlui/handle/915/42994

Michalski, R., Carbonell, J., y Mitchell, T. M. (1983). *Machine learning: An artificial intelligence approach*. Springer.

Mitchell, T. (1997). *Machine learning*. McGraw-Hill.

Moreno, E. (2008). Predicción con series de tiempo y regresión. *Panorama, 2*(4), 36–58. https://revistas.poligran.edu.co/index.php/panorama/article/view/262

Nirmala, H., Archana, J., y Gagana, N. (2024). A survey of integrating deep learning-based missing person detection model into CCTV systems for enhanced identification. *International Journal of Advanced Research in Science, Communication and Technology, 4*(1), 259–268. https://doi.org/10.48175/IJARSCT-15336

Poveda, D., Ramírez, M., Fernández, Y., y García, E. (2024). Implementación del método de optimización de máquinas de soporte vectorial en R. *Tecnura, 28*(80), 99–118. https://doi.org/10.14483/22487638.20326

Quispe, A., Hinojosa, J., Miranda, H., y Sedano, C. (2021). Serie de Redacción Científica: Revisiones sistemáticas. *Revista del Cuerpo Médico del HNAAA, 14*(1), 94–99. https://doi.org/10.35434/rcmhnaaa.2021.141.906

Robledo, C., y Querales, M. (2020). Desaparición de personas en el mundo globalizado: Desafíos desde América Latina. *Íconos, 67*, 1–9. https://www.redalyc.org/journal/509/50963078001/50963078001.pdf

Ruiz, A., Archila, M., y González, J. (2022). Predicting the probability of finding missing older adults based on machine learning. *Journal of Computational Social Science, 5*(2), 1303–1318. https://doi.org/10.1007/s42001-022-00171-x

Russell, S., y Norvig, P. (2020). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

Tumbaco, M., Mendoza, N., Rivera, C., Miranda, K., y Álvarez, G. (2025). Modelos predictivos de inteligencia artificial para prevención de lavado: Análisis comparativo tripartito en la región andina. *Revista Metropolitana de Ciencias Aplicadas, 8*(2), 190–205. https://doi.org/10.62452/f2j9cv82