Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.
ISSN 2707-2207 / ISSN 2707-2215 (en línea), noviembre-diciembre 2024,
Volumen 8, Número 6.

https://doi.org/10.37811/cl_rcm.v8i6

ANÁLISIS DE REVISIÓN SISTEMÁTICA DE LA APLICACIÓN DE ALGORITMOS DE APRENDIZAJE AUTOMÁTICO EN SISTEMAS DE DETECCIÓN DE INTRUSIÓN EN INTERNET DE LAS COSAS PARA CIUDADES INTELIGENTES

ANALYSIS OF SYSTEMATIC REVIEW OF TYHE APPLICATION OF MACHINE LEARNING ALGORITHMS IN INTRUSION DETECTION SYSTEMS IN THE INTERNET OF THINGS FOR SMART CITIES

Luis Uriel Hidalgo Vargas
Universidad Autónoma del Estado de Quintana Roo

José Antonio León Borges
Universidad Autónoma del Estado de Quintana Roo

Julio César Ramírez Pacheco
Universidad Autónoma del Estado de Quintana Roo

Homero Toral Cruz
Universidad Autónoma del Estado de Quintana Roo

Tanya Gabriela Makita Balcorta
Universidad Autónoma del Estado de Quintana Roo

Ismael Osuna Galán
Universidad Autónoma del Estado de Quintana Roo

DOI: https://doi.org/10.37811/cl_rcm.v8i6.15929

Análisis de Revisión Sistemática de la Aplicación de Algoritmos de Aprendizaje Automático en Sistemas de Detección de Intrusión en Internet de las Cosas para Ciudades Inteligentes

Luis Uriel Hidalgo Vargas[1]
1620301@uqroo.edu.mx
https://orcid.org/0009-0006-8698-4929
Universidad Autónoma del Estado de Quintana Roo
México

José Antonio León Borges
jleon@uqroo.edu.mx
https://orcid.org/0000-0002-3992-8696
Universidad Autónoma del Estado de Quintana Roo
México

Julio César Ramírez Pacheco
juliocr@uqroo.edu.mx
https://orcid.org/0000-0002-0397-6063
Universidad Autónoma del Estado de Quintana Roo
México

Homero Toral Cruz
htoral@uqroo.edu.mx
https://orcid.org/0000-0002-4421-3775
Universidad Autónoma del Estado de Quintana Roo
México

Tanya Gabriela Makita Balcorta
t.makita@uqroo.edu.mx
https://orcid.org/0000-0001-6644-8199
Universidad Autónoma del Estado de Quintana Roo
México

Ismael Osuna Galán
ismael.osuna@uqroo.edu.mx
https://orcid.org/0000-0002-3127-7489
Universidad Autónoma del Estado de Quintana Roo
México

RESUMEN

La presente investigación tiene como propósito desarrollar un análisis exhaustivo mediante una revisión sistemática de información, con el fin de responder a la pregunta de investigación: ¿Qué algoritmos implementan el aprendizaje automático en los sistemas de detección de intrusiones en el Internet de las cosas (IoT) en ciudades inteligentes? Para ello, se emplearon estrategias metodológicas basadas en la búsqueda sistemática de artículos científicos extraídos de reconocidas bibliotecas virtuales, utilizando bases de datos como IEEE, ScienceDirect de Elsevier y Springer. Estas fuentes se seleccionaron debido a su alta reputación en la difusión de estudios tecnológicos y científicos de calidad. En una primera etapa, se identificaron y descargaron 125 artículos relacionados con la temática, los cuales pasaron por un proceso riguroso de lectura y evaluación para determinar su relevancia. Este procedimiento incluyó la exclusión de trabajos que no cumplían con los criterios establecidos, garantizando así que solo se consideraran publicaciones directamente vinculadas con el objeto de estudio. El análisis detallado de

[1] Autor principal.
Correspondencia: 1620301@uqroo.edu.mx

los artículos seleccionados permitió identificar los algoritmos más empleados en sistemas de detección de intrusiones que integran aprendizaje automático en el ámbito del IoT, particularmente en el contexto de las ciudades inteligentes. Los resultados obtenidos abarcan el período comprendido entre 2018 y 2023, y reflejan tendencias actuales, fortalezas y limitaciones de los algoritmos utilizados, así como su capacidad para abordar los desafíos de seguridad en entornos interconectados. Este trabajo no solo contribuye a entender el panorama tecnológico en esta área, sino que también proporciona una base sólida para futuras investigaciones orientadas a la optimización y desarrollo de soluciones más eficaces para proteger los sistemas IoT en ciudades inteligentes.

Palabras clave: aprendizaje automático, detección de intrusos, internet de las cosas,ciudades inteligentes

Analysis of Systematic Review of tyhe Application of Machine Learning Algorithms in Intrusion Detection Systems in the Internet of Things for Smart Cities

ABSTRACT

The present research aims to develop an analysis of the systematic review of information to answer the research question: what algorithms implement machine learning in intrusion detection systems in the Internet of Things (IoT) in smart cities?, with the methodological strategies of carrying out a systematic search based on articles extracted from virtual libraries, using databases such as IEEE, ScienceDirect from Elsevier and Springer. 125 downloaded articles are identified, where they are subsequently excluded after reading. The results of the analysis of the articles show the most used algorithms in intrusion detection systems in the field of machine learning, in the period between 2018 and 2023.

***Keywords*:** machine learning, intrusion detection, internet of things, smart cities

Artículo recibido 15 noviembre 2024
Aceptado para publicación: 20 diciembre 2024

INTRODUCCIÓN

En los últimos años, el avance del Internet de las Cosas (IoT) ha impulsado la creación de ciudades inteligentes que operan en tiempo real, mejorando significativamente la calidad de vida en los entornos urbanos (Alrashdi, 2019). Se proyecta que el número de dispositivos IoT a nivel mundial alcanzará los 25 mil millones para el año 2021, con aplicaciones cada vez más frecuentes en sectores como la atención médica, la agricultura, el transporte, los servicios de inteligencia, los hogares y los sistemas urbanos (Ciklabakkal, 2019). El desarrollo de dispositivos inteligentes ha introducido nuevas amenazas que afectan todas las capas de la arquitectura IoT, lo que hace imprescindible diseñar e implementar contramedidas de seguridad adecuadas (Liu, 2019). En este contexto, las ciudades inteligentes emplean diversos procedimientos y métodos para identificar posibles amenazas a la seguridad (Casola, 2019).

Los sistemas de detección de intrusiones (IDS) basados en el aprendizaje automático utilizan técnicas como el entrenamiento, la extracción y selección de características, así como el etiquetado, para clasificar posibles amenazas (Liu, 2019). Sin embargo, los IDS tradicionales no están diseñados para operar eficazmente en redes IoT debido a las limitaciones de recursos y funcionalidad de estos dispositivos. Por ello, el aprendizaje automático se emplea como herramienta clave para identificar y alertar sobre diversos tipos de ataques (Singh, 2021).

El objetivo de este trabajo es analizar a partir de la revisión sistemática sobre la aplicación del aprendizaje automático en sistemas de detección de intrusiones para el Internet de las Cosas en ciudades inteligentes (Hidalgo Vargas, y otros, 2023), con el objetivo de analizar los algoritmos usados en los trabajos, basándose en la literatura publicada entre 2018 y 2023 en tres bibliotecas digitales: IEEE, Springer y Elsevier. Se busca identificar los tipos de algoritmos empleados entre 2018 y 2023 en el ámbito del aprendizaje automático, así como los diferentes tipos de sistemas de detección de intrusiones (IDS) utilizados para proteger las redes contra ataques.

METODOLOGÍA

En análisis se realiza de la revisión sistemática que tuvo como propósito reunir información de bibliotecas virtuales sobre el uso del aprendizaje automático en sistemas de detección de intrusiones para IoT en ciudades inteligentes. Se incluyeron únicamente los artículos que cumplían estrictamente con los criterios establecidos en la investigación. La búsqueda abarca un período de seis años, desde

2018 hasta 2023. Los criterios para la inclusión de artículos fueron los siguientes: Aprendizaje automático, Sistema de detección de intrusióny ciudades inteligentes.

Los criterios de exclusión de artículos son los siguientes:

- Resúmenes
- Comentarios

Los datos se pueden encontrar en diversas bibliotecas virtuales pertenecientes a editoriales reconocidas como IEEE, Springer y Elsevier, las cuales se pueden descargar de las siguientes bases de datos asociadas a cada editorial mencionada:

Se encontraron 125 artículos, de IEEE, 60 de Springer y 54 de Elsevier, fueron excluidos 109 después de revisar título y resumen, y lectura a fondo.

La búsqueda es la siguiente:

- Machine AND learning AND in AND Intrusion AND Detection AND System AND IOT AND networks AND for AND smart AND city' (2018-2023).
- Machine learning AND IDS AND IOT AND SMART CITY NOT survey NOT review.

La técnica para recolección es la siguiente:

- Exportar los artículos encontrados en la búsqueda.
- Cada artículo seleccionado se inspecciona, al no coincidir con el tema se exlcuyen.
- Finalmente, una lectura detallada permite seleccionar y eliminar los artículos

RESULTADOS Y DISCUSIÓN

La Tabla 1 presenta un resumen de las actividades realizadas en cada artículo, incluyendo los tipos de datos trabajados y los objetivos de las investigaciones. Los datos analizados abarcan información de redes, capas de niebla, tráfico de red, paquetes MQTT (Message Queuing Telemetry Transport), vehículos inteligentes, prototipos específicos, recolección de datos de páginas web y otros enfoques orientados a la recopilación de información. Además, se detallan las clasificaciones de ataques, los métodos de detección, las descripciones de diseño de los IDS, y los procesos de extracción y selección de características. También se mencionan algunos prototipos relevantes utilizados en los estudios.

Las actividades del grupo de intervención se enfocan en los programas y sistemas implementados, destacando las funcionalidades de los prototipos. Estos sistemas monitorizan el tráfico de red para

entrenar modelos, realizan procesos de extracción y selección de características, y proponen soluciones para mejorar la seguridad. En la fase de resultados, se evalúa el tráfico de red utilizando varios algoritmos, y se selecciona el más eficiente, siendo Random Forest el preferido en la mayoría de los casos. Asimismo, se mide la precisión y efectividad de los prototipos desarrollados, validando y seleccionando datasets adecuados para la detección de anomalías. Finalmente, se determina la efectividad general de los programas y los logros alcanzados en cada caso.

Tabla 1 Resumen de artículos resultantes de la revisión sistemática

Autor/ año	Muestra	Objetivo	Ejercicios grupo intervención	Resultados
S. Singh (Singh, 2021)**.**	"Red de capa de niebla (FOG) analizada, datos entrenados = 3.768.176, datos probados = 1.925.977" (Singh, 2021).	“Clasificar el tipo de ataque tras su detección” (Singh, 2021).	"Programa para probar datos de red UNSW-NB15" (Singh, 2021).	"Probar el modelo de tráfico de red" (Singh, 2021).
I.Alrashdi (Alrashdi, 2019)**.**	"Tráfico descentralizado en múltiples clasificaciones. Educados = 146.638. Registros totales = 1.422.083" (Alrashdi, 2019).	"Detecta ataques y reduce los falsos positivos con un sistema AD-IOT." (Alrashdi, 2019).	"El sistema AD-IOT monitorea todo el tráfico distribuido en la capa de niebla" (Alrashdi, 2019).	"El estudio evaluó algoritmos de aprendizaje automático para detectar tráfico de red benigno procedente de actividades maliciosas" (Alrashdi, 2019).
Ciklabakkal (Ciklabakkal, 2019)**.**	"Aproximadamente 180 mil paquetes de 100 mil son paquetes MQTT y el resto son principalmente paquetes TCP de paquetes relacionados con MQTT" (Ciklabakkal, 2019).	"Describimos el diseño e implementación de un IDS basado en anomalías para redes IoT con enfoque en ataques MQTT"	“Se crea una tabla que clasifica los ataques benignos, maliciosos y el número total de paquetes” (Ciklabakkal, 2019).	"La detección de intrusiones basada en ML en redes de IoT basadas en MQTT puede lograr resultados impresionantes”

		(Ciklabakkal, 2019).		(Ciklabakkal, 2019).
Liu (Liu, 2019)**.**	“Clasificación y clasificación de conjuntos de datos balanceados y no balanceados, un total de 664.574 probados y 1.531.902 entrenados” (Liu, 2019).	"Se propone un enfoque supervisado con proceso automático de extracción y selección" (Liu, 2019).	"Se realiza la extracción y selección de características. Utiliza cuatro algoritmos para la selección automática: AFES-CART, AFES-RF, AFES-NN, AFES-SVM " (Liu, 2019). .	"El panel izquierdo enumera las funciones seleccionadas. El panel derecho enumera las funciones seleccionadas" (Liu, 2019).
M. Aloqaily (Aloqaily, 2019)**.**	"Número de vehículos utilizados en el estudio = 40" (Aloqaily, 2019).	"El mecanismo de detección de falsos es una solución de monitoreo híbrida que utiliza Deep Belief Network (DBN)" (Aloqaily, 2019).	"El trabajo presentado propuso una solución confiable de búsqueda y selección de servicios para vehículos conectados, donde los vehículos inteligentes (SV) se comunican con los proveedores de servicios a través de cabezales de clúster" (Aloqaily, 2019).	"Realizamos un procedimiento de clasificación de 5 clases para evaluar el rendimiento de D2H-IDS aplica el conjunto de datos NSL-KDD para detectar ataques" (Aloqaily, 2019).
Elsaeidy (Elsaeidy, 2019)**.**	"Si k- significa que es un algoritmo de agrupamiento, entonces k está entre 1 y 10, y el número máximo de iteraciones es 500. En el modelo RBM, el	"Se presenta para crear marcos de seguridad confiables y efectivos para ciudades	“El sistema utilizado para crear la base de datos está relacionado con el sistema inteligente de agua dispositivo experimental,	"Resultados realizados en el conjunto de datos del sistema inteligente de distribución de agua para las

	número de capas está entre 1 y 5, unidades ocultas 50 por cada capa oculta nivel., épocas 500, tasa de aprendizaje 0.001, el tamaño establecido es 100 y los valores de momento inicial y final son 0.5 y 0.9" (Elsaeidy, 2019).	inteligentes" (Elsaeidy, 2019).	rendimiento del método propuesto para detectar diferentes tipos de ataques DDoS" (Elsaeidy, 2019).	diversas configuraciones experimentales analizadas anteriormente" (Elsaeidy, 2019).
Rahman (Rahman, 2021)**.**	"Los autores recopilaron 5.000 páginas web de phishing y 5.000 legítimas para su conjunto de datos basándose en URL y documentos HTML" (Rahman, 2021).	"Marco híbrido basado en la selección de características (HEFS), para obtener características de referencia para la detección de phishing mediante el aprendizaje automático" (Rahman, 2021).	"1. Combinar características 2. Extracción de características" (Rahman, 2021).	"Aplicando SVM a las 20 características de clasificación principales, los resultados numéricos son Acc (99,89%), F1 (99,89%) y MCC (99,77%)." (Rahman, 2021).
Siddiqui (Siddiqui, 2020)**.**	"Basado en transmisiones, el conjunto de datos consta de tráfico malicioso resultante de seis tipos de ataques diferentes: fuerza bruta, Heartbleed, Botnet, DoS y DDoS, ataques web y ataques de	"El método de detección de intrusiones se creó con el objetivo de convertir características basadas en flujo estadístico sin procesar en una representación	"La Tabla 3 de CICIDS2017 muestra la composición del conjunto de datos recopilados durante cinco días (de lunes a viernes) durante cada (nueve) horas de trabajo. NBaIoT: Los conjuntos de	"NBAIoT propone aprender un libro de códigos temporal para cada dispositivo que almacene sus patrones de flujo benignos más importantes" (Siddiqui, 2020).

	infiltración" (Siddiqui, 2020).	más discriminante llamada TempoCode-IoT" (Siddiqui, 2020).	datos" (Siddiqui, 2020).	
Pundir (Pundir, 2021)**.**	“El trabajo se centra en los medios IoT para poder utilizar este prototipo MADP-IIME” (Pundir, 2021).	"Mecanismo de detección de malware en un entorno multimedia industrial habilitado para IoT mediante un enfoque de aprendizaje automático llamado MADP-IIME" (Pundir, 2021).	“Paso 1 Preprocesamiento de datos. Detección de malware en etapa 2” (Pundir, 2021).	"Se introdujo y demostró en la práctica un nuevo sistema de detección de malware que utiliza aprendizaje automático (MADP-IIME) para detectar la presencia de ataques" (Pundir, 2021).
Vigoya Morales (Vigoya Morales, 2023)**.**	“Modelo de red de sensores de temperatura, creado con técnicas estadísticas para añadirlos en dispositivos de baja carga computacional y ligeros dispositivos que simulan los sensores reales” (Vigoya Morales, 2023).	“Presentación, publicación, optimización, análisis y validación de datasets creados para detección de anomalías empleando aprendizaje automático” (Vigoya Morales, 2023).	“Optimización, combinación, modelos y algoritmos de aprendizaje automático de preprocesado de datos, generación y selección de características, discretización y eliminación de información, con la hiperparametrización de modelos y manejo de datos	“Análisis y validación de los datasets para la detección de anomalías empleando aprendizaje automático” (Vigoya Morales, 2023).

			desbalanceados.” (Vigoya Morales, 2023).	
Park (Park, 2018)	"La recopilación de datos se clasifica en datos del sistema operativo y datos de las instalaciones" (Park, 2018).	"Este estudio estableció una arquitectura de evaluación de intrusiones basada en aprendizaje automático, calculó la tasa de detección de anomalías y la eficiencia, y comparó los efectos" (Park, 2018).	"Fase de recopilación y análisis de datos. Fase de construcción y discusión del modelo. Está optimizado mediante aprendizaje de patrones y aprendizaje iterativo utilizando aprendizaje estándar no supervisado, como agrupación en clústeres y codificador automático.” (Park, 2018).	"La reducción se logró mediante la construcción de un sistema de detección de intrusos basado en el aprendizaje automático. El sistema propuesto mostró 33.33 logros de procesos y más del 29.29 por ciento de detección de anomalías” (Park, 2018).

Fuente: (Hidalgo Vargas, y otros, 2023)

Tabla 2 Comparativa de los artículos con el tipo de IDS

Autores / Artículos	**Nombre del prototipo**	**HIDS (Host Intrusion Detection System)**	**NIDS (Network Intrusion Detection System)**	**NBA (Network Behavior Analysis IDS)**	**WIDS (wireless Intrusion Detection System)**	**MQTT (Message Queuing Telemetry Transport)**	**Nube / niebla**
I.Alrashdi (Alrashdi, 2019)	AD-loT		✓				
Ciklabakkal (Ciklabakkal, 2019)	Artemis					✓	✓

Liu (Liu, 2019)	Método AFES				✓		
S. Singh (Singh, 2021)	UNSW-NB15	✓	✓				
M. Aloqaily (Aloqaily, 2019)	D2H-IDS		✓		✓		
Elsaeidy (Elsaeidy, 2019)	Modelo RBM		✓	✓			
Rahman (Rahman, 2021)	SVM, Aegean Wi-Fi Intrusion Dataset		✓		✓		
Siddiqui (Siddiqui, 2020)	TempoCode-IoT		✓	✓			
Pundir (Pundir, 2021)	MADP-IIME,		✓				✓
Vigoya Morales (Vigoya Morales, 2023)	DAD, CIDAD					✓	
Park (Park, 2018)	AMI, DNN		✓				

Fuente: (Hidalgo Vargas, y otros, 2023)

La Tabla 2 se observa la clasificación del tipo de IDS de los artículos de la revisión sistemática según (Aguilar, 2018), (Ciklabakkal, 2019), (Liu, 2019), (Singh, 2021), (Aloqaily, 2019), (Elsaeidy, 2019), (Rahman, 2021), (Siddiqui, 2020), (Pundir, 2021), (Vigoya Morales, 2023), (Park, 2018), (Martinez,

2017).

En la

Tabla **3** muestra una comparación sobre los algoritmos usados en los últimos 6 años 2018-2023 de aprendizaje automático en los sistemas de detección de intrusiones. Los algoritmos listados son los siguientes:

- Naive Bayes: Un modelo probabilístico utilizado en clasificación, basado en el teorema de Bayes.
- Decision Tree (RF): Árboles de decisión, incluidos en la variante Random Forest (RF), que combina múltiples árboles para mejorar la precisión del modelo.
- Neural Network: Redes neuronales artificiales utilizadas para clasificación o regresión, que imitan el funcionamiento del cerebro.
- Deep Learning: Un enfoque de redes neuronales con múltiples capas, útil para manejar grandes volúmenes de datos complejos.
- Auto Encoder: Red neuronal utilizada principalmente para la reducción de dimensionalidad y la detección de anomalías.
- Active Learning: Un enfoque en el que el modelo selecciona activamente los ejemplos más informativos para aprender.
- Isolation Forest: Un algoritmo de detección de anomalías que aísla las observaciones.
- SVM (Support Vector Machine): Un modelo de clasificación que busca la mejor separación entre clases.
- LR (Logistic Regression): Un modelo de clasificación basado en una función logística, utilizado en problemas binarios.
- Adaboost: Un algoritmo de aprendizaje en conjunto que ajusta los pesos de los clasificadores según su rendimiento.

Se puede observar como el algoritmo Decision Tree es el más utilizado en la actualidad y el algoritmo Adaboost es el menos utilizado, junto con Autoencoder, Isolation Forest y Active Learning se utilizan en menor medida, pero pueden ofrecer ventajas en tareas específicas de detección de ataques. Con una "X" señala el algoritmo implementado en cada artículo para el prototipo. Naive Bayes es el segundo algoritmo más incurrido en los estudios, con la mayoría de los autores considerando su efectividad para

la detección de ataques. Decision Tree Random Forest (RF) es muy popular, siendo utilizado por varios autores debido a su simplicidad y capacidad de interpretación. Neural Networks y Deep Learning son seleccionados por autores que probablemente están trabajando con conjuntos de datos más grandes y complejos. Isolation Forest aparece en trabajos que se centran en la detección de anomalías. Active Learning, SVM, RF y Adaboost son los algoritmos explorados por (Ciklabakkal, 2019) y (Vigoya Morales, 2023) de los once trabajos resultantes.

Tabla 3 Algoritmos usados para la detección de ataques

Autores / artículos	NB	Decision tree (RF)	NN	DL	Auto encoder	AL	Isolation forest	SVM	LR	Adaboost
S. Singh (Singh, 2021)		X		X						
I. Alrashdi (Alrashdi, 2019)			X							
Ciklabakkal (Ciklabakkal, 2019)		X			X	X	X	X		
Liu (Liu, 2019)		X	X							
M. Aloqaily (Aloqaily, 2019)		X								
Elsaeidy (Elsaeidy, 2019)				X						
Rahman (Rahman, 2021)	X	X								

Siddiqui (Siddiqui, 2020)		X							
Pundir (Pundir, 2021)	X	X	X					X	
Vigoya Morales (Vigoya Morales, 2023)	X	X					X	X	X
Park (Park, 2018)				X					

Nota: NB(Naïve Bayes), RF (Random Forest), NN (Neural Network), AL (Active Learning), DL (Deep Learning), SVM (Support Vector Machine), LR (Linear Regression).

La Tabla 4 presenta una lista de autores junto con los artículos correspondientes y sus áreas de enfoque, relacionadas con los componentes de las ciudades inteligentes (smart cities) a partir de (Sikora Fernández, 2017). Cada área de la ciudad inteligente con una "X" si el autor aborda ese tema en su trabajo. A continuación, se describe el contenido de la tabla en función de los temas abordados por cada autor:

- Economía (smart economy): Incluye artículos que tratan sobre la economía inteligente, que puede implicar el uso de tecnología y datos para mejorar la eficiencia económica en las ciudades.
- Transporte y comunicación (smart mobility): Se refiere a los estudios sobre cómo mejorar los sistemas de transporte y las comunicaciones en las ciudades mediante tecnologías inteligentes.
- Medioambiente (smart environment): Trata sobre cómo las tecnologías pueden contribuir a la sostenibilidad y la gestión eficiente del medioambiente en las ciudades.
- Personas (smart people): Examina la participación y el bienestar de las personas en un entorno urbano inteligente, promoviendo la educación, la salud y la interacción social.
- Calidad de vida (smart living): Relacionado con la mejora de la calidad de vida de los residentes mediante el uso de tecnologías inteligentes para una vida más cómoda y eficiente.
- Gestión y administración inteligente (smart governance): Hace referencia al uso de la tecnología

para mejorar la administración pública y la toma de decisiones en la gestión de la ciudad.

Tabla 4 Resultados de los artículos en tipos de ciudades inteligentes / IoT

Autores / Artículos	Economía (smart economy)	Transporte y comunicación (smart mobility)	Medioambiente (smart environment)	Personas (smart people)	Calidad de vida (smart living)	Gestión y administración inteligente (smart governance)
I.Alrashdi (Alrashdi, 2019)	X	X	X	X	X	X
Ciklabakkal (Ciklabakkal, 2019)	X	X			X	X
Liu (Liu, 2019)	X	X			X	X
S. Singh (Singh, 2021)	X			X	X	X
M. Aloqaily (Aloqaily, 2019)		X		X	X	
Elsaeidy (Elsaeidy, 2019)	X	X		X	X	X
Rahman (Rahman, 2021)	X	X		X	X	
Siddiqui (Siddiqui, 2020)		X		X	X	X
Pundir (Pundir, 2021)	X		X	X	X	
Vigoya Morales (Vigoya Morales, 2023)	X	X		X	X	
Park (Park, 2018)	X		X	X	X	

Fuente: (Hidalgo Vargas, y otros, 2023)

El aprendizaje automático es un campo con amplias posibilidades de desarrollo y una variedad de herramientas que permiten abordar problemas complejos mediante algoritmos que ofrecen soluciones innovadoras y versátiles. Este enfoque resulta útil para resolver problemas que, en principio, podrían parecer insolubles. Su aplicación se estructura a través de procesos y etapas definidas por cada autor en función del prototipo y la situación específica en la que se implemente. Además, se incorpora el uso de Sistemas de Detección de Intrusiones (IDS), que cumplen funciones variadas según el contexto. En este

proceso, el aprendizaje automático recopila y analiza información de la red, entrenándose de forma autónoma o con intervención humana, para posteriormente transmitir los resultados al IDS, encargado de garantizar la seguridad de la red mediante el uso de diferentes tipos de IDS adaptados a escenarios específicos en ciudades inteligentes. A partir de los artículos revisados, se concluye que, entre los cuatro tipos de aprendizaje existentes, los más utilizados en el período 2018-2023 han sido el aprendizaje supervisado, no supervisado y semi-supervisado. Cada uno de estos enfoques cuenta con una variedad de algoritmos aplicados a proyectos específicos, siendo los más destacados el algoritmo Random Forest (bosque aleatorio), seguido de las redes neuronales y el aprendizaje profundo (Deep Learning).

Los algoritmos empleados en aprendizaje automático, según los artículos revisados, incluyen: Naive Bayes, Árbol de Decisión (Random Forest), Redes Neuronales, Aprendizaje Profundo, Codificadores Automáticos, Aprendizaje Activo, Bosques de Aislamiento, Máquinas de Vectores de Soporte, Aprendizaje en Selva y Regresión Logística. Estos algoritmos se utilizan para abordar problemas mediante datasets ya entrenados o boost datasets, maximizando la eficiencia de los resultados. Existen diferentes tipos de IDS diseñados para áreas específicas, dependiendo del problema a resolver. Entre ellos se encuentran: WIDS (IDS basados en host), HIDS (IDS basados en red), NIDS (IDS basados en nodos), NBA (Análisis de comportamiento de red) y protocolos como MQTT en entornos Cloud o Fog, los cuales incrementan las soluciones disponibles frente a los ataques de red.

Por otro lado, las ciudades inteligentes que integran tecnologías IoT se encuentran en diferentes etapas de desarrollo, y no todas están completamente actualizadas o automatizadas. Aquellas que ya han adoptado estos sistemas, se han identificado elementos clave que deben ser considerados para su gestión inteligente: economía, transporte y comunicación, medioambiente, personas, calidad de vida, y administración eficiente. Estos factores determinarán las áreas específicas donde se priorizará la mejora de la seguridad en el contexto de las ciudades inteligentes.

CONCLUSIONES

La presente investigación tuvo como objetivo desarrollar un análisis de la revisión sistemática para responder a la pregunta de investigación: ¿qué algoritmos implementan el aprendizaje automático en los sistemas de detección de intrusiones en el Internet de las cosas (IoT) en ciudades inteligentes? los resultados en el análisis de la revisión sistemática ayuda a comprender la respuesta a la pregunta

previamente planteada donde se encuentran algoritmos de aprendizaje automático Naive Bayes, Árbol de Decisión (Random Forest), Redes Neuronales, Aprendizaje Profundo, Codificadores Automáticos, Aprendizaje Activo, Bosques de Aislamiento, Máquinas de Vectores de Soporte, Aprendizaje en Selva y Regresión Logística. La literatura en el período comprendido entre 2018 y 2023 muestra que el algoritmo más usado fue Árbol de Decisión (Random Forest) en los sistemas de detección de intrusiones en el ámbito del aprendizaje automático, sin embargo, se observa un creciente interés por técnicas más avanzadas, como Deep Learning y Neural Networks, lo que refleja la tendencia hacia modelos más complejos en la detección de ataques.

REFERENCIAS BIBLIGRAFICAS

Aguilar, R. M. (2018). Aprendizaje automático en la identificación de sistemas. Un caso de estudio en la predicción de la generación eléctrica de un parque eólico. *Revista Iberoamericana de Automática e Informática.*, 114-127.

Aloqaily, M. (2019). An intrusion detection system for connected vehicles in smart cities. *ELSEVIER*, 27.

Alrashdi, I. (2019). AD-IoT: Anomaly Detection of IoT Cyberattacks in Smart City Using Machine Learning. *IEEE*, 305-310.

Casola, V. (2019). A security monitoring system for internet of things. *ELSEVIER*, 20.

Ciklabakkal, E. (2019). ARTEMIS: An Intrusion Detection System for MQTT Attacks in Internet of Things. *IEEE*, 3.

Elsaeidy, A. (2019). Intrusion detection in smart cities using Restricted Boltzmann Machines. *ELSEVIER*, 76-86.

Hidalgo Vargas, L., Ramírez Pacheco, J., León Borges, J., Osuna Galan, I., Toral Cruz, H., & Makita Balcorta, T. (11 de 2023). Revisión sistemática de la aplicación de aprendizaje automático en sistemas de detección de intrusión en internet de las cosas de ciudades inteligentes. Quintana Roo, México: UQRoo.

Li, Z. D. (2001). Intrusion Detection System. *Computer Engineering*, 7-9.

Liu, J. (2019). Automatic Feature Extraction and Selection For Machine Learning Based Intrusion Detection. *IEEE*, 1400-1405.

Martinez, G. R. (2017). Sistema de detección de intrusos en redes corporativas. *Scientia et technica*, 60-68.

Park, S.-T. (2018). A study on smart factory-based ambient intelligence context-aware intrusion detection system using machine learning. *Springer*, 1-8.

Pundir, S. (2021). MADP-IIME: malware attack detection protocol in IoT-enabled industrial multimedia environment using machine learning approach. *Springer*, 1-13.

Rahman, M. A. (2021). Effective combining of feature selection techniques for machine learning-enabled IoT intrusion detection. *Springer*, 31381–31399.

Siddiqui, A. J. (2020). TempoCode-IoT: temporal codebook-based encoding of flow features for intrusion detection in Internet of Things. *Springer* , 17-35.

Sikora Fernández, D. (2017). Factores de desarrollo de las ciudades inteligentes. *Revista Universitaria de Geografía, 26*, 135-152.

Singh, S. (2021). MCIDS-Multi Classifier Intrusion Detection system for IoT Cyber Attack using Deep Learning algorithm. *IEEE* , 354-360.

Tux. (06 de febrero de 2012). *Qué son y para qué sirven*. Obtenido de http://www.g3ekarmy.com/que-son-y-para-que-sirven-los-firewalls/

Vigoya Morales, L. (2023). Aplicación de algoritmos de aprendizaje automático para la detección de anomalías de tráfico en entornos IoT. *Universidade da Coruña.*, 1-142.