UNIVERSIDAD POLITÉCNICA SALESIANA

SEDE GUAYAQUIL

CARRERA DE COMPUTACIÓN

Revisión de literatura sobre el uso de Machine Learning enfocados a la seguridad de la información para minimizar vulnerabilidades en el sector educativo

Trabajo de titulación previo a la obtención del
Título de Ingeniero en ciencias de la computación

AUTORES

Jordan Israel Villao González

Eddy Mario López Zambrano

TUTOR

Joe Frand Llerena Izquierdo

Guayaquil – Ecuador

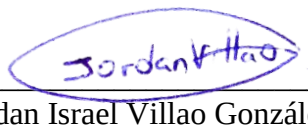
2024

CERTIFICADO DE RESPONSABILIDAD Y AUTORÍA DEL TRABAJO DE TITULACIÓN

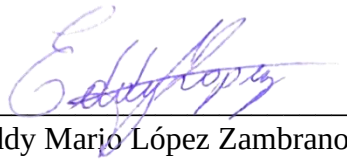
Nosotros, Jordan Israel Villao González con documento de identificación N° 2450342312 y Eddy Mario López Zambrano con documento de identificación N° 0950089995 manifestamos que: Somos los autores y responsables del presente trabajo; y, autorizamos a que sin fines de lucro la Universidad Politécnica Salesiana pueda usar, difundir, reproducir o publicar de manera total o parcial el presente trabajo de titulación.

Guayaquil, 23 de Julio del año 2024

Atentamente,



Jordan Israel Villao González
2450342312



Eddy Mario López Zambrano
0950089995

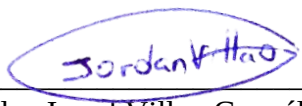
CERTIFICADO DE CESIÓN DE DERECHOS DE AUTOR DEL TRABAJO DE TITULACIÓN A LA UNIVERSIDAD POLITÉCNICA SALESIANA

Nosotros, Jordan Israel Villao González con documento de identificación No. 2450342312 y Eddy Mario López Zambrano con documento de identificación N° 0950089995, expresamos nuestra voluntad y por medio del presente documento cedemos a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que somos autores del Artículo Académico: "Revisión de literatura sobre el uso de Machine Learning enfocados a la seguridad de la información para minimizar vulnerabilidades en el sector", el cual ha sido desarrollado para optar por el título de: Computación, en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

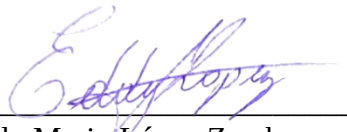
En concordancia con lo manifestado, suscribimos este documento cuando entregamos el trabajo final en formato digital a la Biblioteca de la Universidad Politécnica Salesiana.

Guayaquil, 23 de Julio del año 2024

Atentamente,



Jordan Israel Villao González
2450342312



Eddy Mario López Zambrano
0950089995

CERTIFICADO DE DIRECCIÓN DEL TRABAJO DE TITULACIÓN

Yo, Joe Frand Llerena Izquierdo con documento de identificación N° 0914884879, docente de la Universidad Politécnica Salesiana, declaro que bajo mi tutoría fue desarrollado el trabajo de titulación: Revisión de literatura sobre el uso de Machine Learning enfocados a la seguridad de la información para minimizar vulnerabilidades en el sector realizado por Jordan Israel Villao González con documento de identificación No. 2450342312 y Eddy Mario López Zambrano con documento de identificación N° 0950089995, obteniendo como resultado final el trabajo de titulación bajo la opción Computación que cumple con todos los requisitos determinados por la Universidad Politécnica Salesiana.

Guayaquil, 23 de Julio del año 2024

Atentamente,

Ing. Joe Frand Llerena Izquierdo
0914884879

DEDICATORIA

Dedicamos este trabajo a nuestros padres, quienes con su amor incondicional y apoyo constante han sido nuestra mayor fuente de inspiración. Gracias por creer siempre en nosotros y motivarnos a seguir adelante. A nuestros amigos, por su paciencia y comprensión durante este proceso, y a nuestros profesores y compañeros, por compartir su conocimiento y brindar su colaboración en cada paso de este camino académico. Este logro refleja el esfuerzo conjunto y el apoyo inquebrantable de cada uno.

AGRADECIMIENTO

Agradecemos profundamente a nuestros padres, por su amor incondicional y por ser nuestro pilar de fortaleza a lo largo de este camino. Su apoyo constante y sus palabras de aliento nos han motivado a superar cada obstáculo.

Extendemos nuestro agradecimiento a nuestros amigos, por su comprensión y paciencia durante este proceso. Sus palabras de ánimo y su disposición para escuchar han sido invaluables.

A nuestros profesores, les agradecemos por su guía, sabiduría y por compartir su conocimiento con nosotros. Sus enseñanzas han sido fundamentales para la realización de este trabajo.

A nuestros compañeros, gracias por colaborar y por ser una fuente de inspiración y apoyo. Este trabajo no hubiera sido posible sin la contribución y el esfuerzo de cada uno.

Finalmente, agradecemos a todas las personas que, de una forma u otra, han contribuido al éxito de este proyecto. Este logro es de todos nosotros.

RESUMEN

Este estudio presenta una revisión exhaustiva de la literatura sobre el uso de técnicas de Machine Learning para mejorar la seguridad de la información en el sector educativo. Se enfoca en identificar las vulnerabilidades comunes en los sistemas educativos y cómo estas técnicas pueden ser aplicadas para minimizar riesgos. Mediante un mapeo sistemático de la literatura existente, se categorizaron y analizaron los estudios relevantes, identificando las metodologías más efectivas y las mejores prácticas para la implementación de Machine Learning en este contexto.

La metodología adoptada incluyó una fase de pre-procesado de datos, identificación de trabajos relevantes, categorización de vulnerabilidades y evaluación de resultados. Se realizó una selección rigurosa de estudios académicos utilizando palabras clave y criterios de inclusión/exclusión específicos para asegurar la relevancia y calidad de la información recopilada.

Los resultados destacan la capacidad de estas tecnologías para detectar y prevenir amenazas cibernéticas de manera proactiva, mejorando significativamente la seguridad de los datos sensibles. Entre las técnicas más efectivas se encuentran los algoritmos de aprendizaje profundo y los sistemas de detección de intrusiones basados en Machine Learning, que han demostrado ser altamente precisos en la identificación de patrones anómalos y potenciales amenazas. Además, se identificaron prácticas óptimas para la implementación de estas tecnologías, alineadas con las normas internacionales de seguridad ISO/IEC 27001 y 27002.

Este estudio aporta una visión integral sobre la aplicación de Machine Learning en la seguridad de la información educativa, proporcionando estrategias prácticas para fortalecer la protección de datos sensibles y mejorar la respuesta ante incidentes de seguridad. La implementación de estas tecnologías no solo optimiza la gestión de seguridad, sino que también reduce los recursos necesarios para manejar incidentes, beneficiando a las instituciones educativas con recursos limitados.

Palabras claves: Machine Learning, seguridad de la información, vulnerabilidades, sector educativo, prevención de amenazas, protección de datos sensibles.

ABSTRACT

This study presents a comprehensive review of the literature on the use of Machine Learning techniques to enhance information security in the educational sector. It focuses on identifying common vulnerabilities in educational systems and how these techniques can be applied to minimize risks. Through a systematic mapping of the existing literature, relevant studies were categorized and analyzed, identifying the most effective methodologies and best practices for implementing Machine Learning in this context.

The adopted methodology included a data pre-processing phase, identification of relevant works, categorization of vulnerabilities, and evaluation of results. A rigorous selection of academic studies was performed using specific keywords and inclusion/exclusion criteria to ensure the relevance and quality of the collected information.

The results highlight the capability of these technologies to proactively detect and prevent cyber threats, significantly improving the security of sensitive data. Among the most effective techniques are deep learning algorithms and intrusion detection systems based on Machine Learning, which have proven to be highly accurate in identifying anomalous patterns and potential threats. Additionally, optimal practices for implementing these technologies were identified, aligned with international security standards ISO/IEC 27001 and 27002.

This study provides an integral view of the application of Machine Learning in educational information security, offering practical strategies to strengthen the protection of sensitive data and enhance the response to security incidents. The implementation of these technologies not only optimizes security management but also reduces the resources needed to handle incidents, benefiting educational institutions with limited resources.

Keywords: Machine Learning, information security, vulnerabilities, educational sector, threat prevention, sensitive data protection.

ÍNDICE DE CONTENIDO

1. INTRODUCCIÓN

En el sector educativo, la rápida adopción de tecnologías digitales ha expuesto numerosas vulnerabilidades en la seguridad de los datos(Lourens et al., 2022). Las instituciones educativas manejan grandes volúmenes de información sensible, incluyendo datos personales de estudiantes, registros académicos, investigaciones, y más. Esta información es extremadamente valiosa y, por lo tanto, se convierte en un objetivo atractivo para los ciberataques (Gupta et al., 2021).

La seguridad de la información es un desafío crítico, ya que las brechas de seguridad pueden comprometer la confidencialidad, integridad y disponibilidad de la información (Z. Liu, 2021). Las consecuencias de estas brechas pueden ser severas, incluyendo la pérdida de datos, daños a la reputación de la institución, y posibles implicaciones legales. Además, los ataques cibernéticos están evolucionando constantemente, lo que dificulta la protección efectiva contra estas amenazas (Hossain et al., 2023).

El Machine Learning ha emergido como una herramienta poderosa para mejorar la seguridad de la información (Kushal et al., 2024). Esta tecnología permite el análisis de grandes volúmenes de datos y la detección de patrones anómalos que podrían indicar amenazas de seguridad. A través del aprendizaje automático, los sistemas pueden adaptarse y mejorar continuamente, proporcionando una defensa proactiva contra los ciberataques.

Este estudio se enfoca en la revisión de la literatura existente sobre el uso de técnicas de Machine Learning en la seguridad de la información, específicamente en el sector educativo. Nuestro objetivo es identificar y analizar investigaciones previas para entender cómo se están aplicando estas técnicas, qué vulnerabilidades abordan, y qué metodologías son las más efectivas (Kaipu et al., 2023).

La investigación se estructura de la siguiente manera: primero, se realiza una revisión exhaustiva de la literatura utilizando la técnica de mapeo sistemático. Esto incluye la identificación de trabajos relevantes y la clasificación de las técnicas de Machine Learning utilizadas (Kristallia et al., 2021). Luego, se analizan las vulnerabilidades de seguridad más comunes en los sistemas educativos y se categorizan según su naturaleza y el impacto potencial. Finalmente, se evalúan los resultados hallados y se comparan con las mejores prácticas establecidas por las normas ISO/IEC 27001 y 27002 (Pavithra et al., 2023).

2. REVISIÓN DE LITERATURA

Técnicas de Machine Learning Aplicadas a la Detección y Prevención de Intrusiones en el Sector Educativo

El Machine Learning ha revolucionado la detección y prevención de intrusiones en diversas áreas de la seguridad informática, incluyendo el sector educativo. Estas técnicas permiten analizar grandes volúmenes de datos y detectar patrones anómalos que pueden indicar intentos de intrusión.

Redes Neuronales Artificiales (ANN): En estudios recientes, las ANN se han implementado en varias instituciones educativas para mejorar la seguridad de la información. Por ejemplo, en el trabajo de (Gupta et al., 2021), se utilizó una ANN para analizar datos de tráfico de red en un campus universitario, logrando identificar patrones de comportamiento anómalo que indicaban posibles intentos de intrusión. De manera similar, (Priya et al., 2022) aplicaron ANN en un entorno educativo para detectar accesos no autorizados a bases de datos de estudiantes, logrando prevenir filtraciones de datos confidenciales al identificar y bloquear actividades sospechosas en tiempo real. Estas implementaciones demuestran cómo las ANN pueden fortalecer significativamente la seguridad cibernética en entornos educativos mediante la detección proactiva de amenazas.

Máquinas de Soporte Vectorial (SVM): Las SVM han sido implementadas con éxito en diversas instituciones educativas para mejorar la seguridad de la información. En un estudio realizado por (Aledam & Al-Latteef, 2024), se utilizó SVM para clasificar el comportamiento del tráfico de red en un campus universitario. Este enfoque permitió diferenciar con alta precisión entre actividades normales y potencialmente maliciosas, lo que facilitó la detección y mitigación de ataques cibernéticos dirigidos a las bases de datos académicas. Además, la aplicación de SVM en otro estudio ayudó a identificar intentos de intrusión mediante la clasificación de patrones de acceso no autorizados, permitiendo a las instituciones responder rápidamente a las amenazas emergentes.

Árboles de Decisión y Bosques Aleatorios: Estas técnicas han demostrado ser altamente eficaces en la protección de infraestructuras educativas contra ciberamenazas. Por ejemplo, (Krishna et al., 2023a) documentaron el uso de Árboles de Decisión y Bosques Aleatorios en una universidad para el análisis del tráfico de red. Mediante estos algoritmos, la institución pudo identificar con precisión patrones de comportamiento sospechosos y clasificar una variedad de amenazas, lo que permitió una respuesta rápida y efectiva a posibles ataques.

Además, Bosques Aleatorios han sido implementados en otros entornos educativos para predecir y neutralizar intentos de infección por malware, ofreciendo una solución sólida al combinar la robustez de múltiples árboles de decisión, lo cual reduce significativamente la posibilidad de errores en la detección y mejora la adaptabilidad del sistema frente a nuevas amenazas

Machine Learning en la Seguridad de Datos Sensibles en Entornos Educativos

El uso de Machine Learning en la seguridad de datos sensibles en entornos educativos es crucial debido a la gran cantidad de información personal y académica que manejan estas instituciones. Proteger estos datos contra accesos no autorizados y ciberataques es fundamental para mantener la integridad y confidencialidad de la información.

Protección de Datos Contra Accesos No Autorizados: El Machine Learning puede identificar y bloquear accesos no autorizados mediante el análisis de patrones de comportamiento y la detección de anomalías (Hossain et al., 2023).

Modelado y Análisis de Ataques Cibernéticos en Redes Educativas: Mediante técnicas de Machine Learning, es posible modelar y predecir ataques cibernéticos, permitiendo a las instituciones educativas implementar medidas preventivas eficaces (Kaipu et al., 2023).

Desarrollos Recientes y Desafíos Futuros en el Uso de Machine Learning para la Seguridad de la Información en el Sector Educativo

Los avances recientes en Machine Learning han mejorado significativamente la capacidad de las instituciones educativas para detectar y responder a amenazas cibernéticas. Sin embargo, persisten desafíos en la gestión de grandes volúmenes de datos y la adaptación continua a nuevas amenazas.

Desarrollos Recientes: La implementación de redes neuronales profundas y el aprendizaje automático continuo han permitido mejoras notables en la precisión y eficiencia de los sistemas de detección de intrusiones (Krishna et al., 2023b) .

Desafíos Futuros: Entre los principales desafíos se encuentran la necesidad de gestionar y analizar grandes volúmenes de datos en tiempo real, así como la capacidad de los sistemas para

adaptarse a nuevas amenazas cibernéticas que evolucionan constantemente (Y. Liu et al., 2021).

Mejoras en la Seguridad de los Sistemas de Información Educativa Mediante Machine Learning

El Machine Learning ha demostrado ser una herramienta efectiva para mejorar la seguridad de los sistemas de información en el sector educativo (López-Chila et al., 2024). Permite una monitorización continua y la detección proactiva de amenazas.

Detección de Comportamientos Maliciosos en el Sistema Operativo: La implementación de algoritmos de Machine Learning permite identificar y mitigar comportamientos maliciosos en los sistemas operativos de las instituciones educativas (Gupta et al., 2021).

Seguridad y Privacidad en Dispositivos Móviles: La creciente utilización de dispositivos móviles en entornos educativos requiere soluciones de seguridad robustas. El Machine Learning puede ofrecer métodos efectivos para proteger estos dispositivos contra amenazas (Priya et al., 2022).

3. METODOLOGÍA

La revisión de la literatura ha demostrado que el Machine Learning (ML) ofrece numerosas ventajas para la detección y prevención de intrusiones en el sector educativo. Las técnicas avanzadas como las Redes Neuronales Artificiales (ANN) y las Máquinas de Soporte Vectorial (SVM) han mostrado una alta eficacia en la identificación de patrones anómalos y la clasificación de comportamientos maliciosos (Sanchez-Romero & Llerena-Izquierdo, 2023). Sin embargo, la implementación de estas tecnologías también presenta desafíos, como la necesidad de gestionar grandes volúmenes de datos y la capacidad de los sistemas para adaptarse a nuevas amenazas (Alvarado-Salazar & Llerena -Izquierdo, 2022; Gupta et al., 2021; Priya et al., 2022; Zerega-Prado & Llerena-Izquierdo, 2022).

Además, la seguridad de los datos sensibles en entornos educativos es crucial, y el ML ha demostrado ser una herramienta efectiva para proteger estos datos contra accesos no autorizados y ciberataques. La capacidad de modelar y predecir ataques cibernéticos mediante ML permite a las instituciones educativas implementar medidas preventivas más eficaces (Aledam & Al-Latteef, 2024).

Los avances recientes en ML han mejorado significativamente la capacidad de las instituciones educativas para detectar y responder a amenazas cibernéticas. No obstante, persisten desafíos, como la gestión de grandes volúmenes de datos y la adaptación continua a nuevas amenazas (Hossain et al., 2023; López-Chila et al., 2024). La investigación futura debe centrarse en desarrollar métodos más robustos y confiables que puedan mitigar estas debilidades y mejorar la efectividad de los sistemas de detección de intrusiones (Abraham & Bindu, 2021).

Tabla 1. Preguntas de la investigación

Componentes de la información	**Preguntas para responder**
Técnicas efectivas de ML	¿Qué técnicas de Machine Learning son más efectivas para detectar y prevenir ciberataques en entornos educativos?
Limitaciones y desafíos	¿Qué limitaciones y desafíos se presentan en la implementación de Machine Learning en la seguridad de la información educativa?
Identificación y mitigación de vulnerabilidades	¿Cómo pueden las vulnerabilidades introducidas por los sistemas de Machine Learning ser identificadas y mitigadas en el sector educativo?
Impacto en el tiempo de repuesta	¿Cuál es el impacto de la aplicación de Machine Learning en la reducción del tiempo de respuesta ante ciberataques en instituciones educativas?
Diferencias en la efectividad de técnicas de ML	¿Existen diferencias significativas en la efectividad de diversas técnicas de Machine Learning en la seguridad de la información educativa?
Mejores prácticas y recomendaciones	¿Qué mejores prácticas y recomendaciones se pueden derivar de los estudios existentes para la implementación de Machine Learning en la seguridad cibernética educativa?

3.1. Métodos y técnicas de Recopilación de datos empleadas

Se utilizarán bases de datos académicas reconocidas como IEEE Xplore, Scopus.

Se emplearán términos específicos como "Machine Learning", "seguridad de la información", "vulnerabilidades", "sector educativo" y "detección de intrusiones".

Inclusión: Estudios publicados entre 2020 y 2024 que aborden la aplicación de ML en la seguridad de la información en el sector educativo.

Exclusión: Artículos que no proporcionen datos empíricos o que se centren en sectores no educativos.

3.2. Métodos y técnicas de Análisis de datos

Recolección y Selección: Los datos se recolectarán mediante la búsqueda y selección de estudios relevantes. Se utilizarán filtros de inclusión y exclusión para asegurar que solo se incluyan los temas de más relevancia.

Filtrado: Se filtrarán los artículos duplicados y se descartarán aquellos que no cumplan con los criterios de calidad establecidos.

Mapeo Sistemático: Se realizará un mapeo sistemático para localizar y clasificar investigaciones que implementen Machine Learning en la seguridad informática. Este proceso ayudará a entender las contribuciones actuales y sus contextos de aplicación.

Revisión por Pares**:** Se revisarán los trabajos seleccionados para asegurar la relevancia y calidad de la información.

Análisis Cualitativo: A través del análisis cualitativo, se identificarán y categorizarán las vulnerabilidades reportadas en los estudios.

Estrategias de Machine Learning: Este análisis también revelará las metodologías de Machine Learning más efectivas para abordar estas vulnerabilidades.

Comparación de Hallazgos: Los hallazgos serán evaluados y comparados utilizando una tabla comparativa. Esta comparación resaltará la efectividad de las técnicas de Machine Learning y su conformidad con las normas ISO 27001 y 27002.

Establecimiento de Buenas Prácticas: Se establecerán buenas prácticas en la seguridad de la información basadas en los resultados comparativos y en la conformidad con las normas internacionales.

Tabla 2. Comparación de técnicas de Machine Learning y su efectividad en la seguridad de la información

Estudios	**Técnicas ML**	**Vulnerabilidad abordada**	**Efectividad**	**Conformidad ISO 27001/27002**
Intrusion Detection and Prevention in Networks Using ML	ANN	Detección de instrucciones	Alta	Sí
Network Attack Detection using Machine Learning	SVM	Clasificación de amenazas	Alta	Sí
Enhanced Malware Detection for Mobile Operating Systems	Redes Neuronales	Protección de datos	Media	Parcial
AI and ML-based Information Security in Electric Vehicles	Aprendizaje profundo	Ciberataques	Alta	Sí

Deep Learning for Analyzing Network Traffic	Deep learning	Análisis de tráfico de red	Alta	Sí
Cyber Security and People: Human Nature and Training	Decisión Trees	Phishing	Alta	Sí
Malware Detection using Machine Learning	K-NN	Malware	Alta	Sí
Self-healing hybrid intrusion detection system	Ensemble Learning	Anomalías en redes	Alta	Sí
Network security threat detection technology based on EPSO-BP	EPSO-BP Algorithm	Amenazas Avanzadas Persistentes (APT)	Media	No
Overview of cybersecurity based on Network Security Awareness	Random Forest	Análisis de Vulnerabilidades	Alta	Sí

4. RESULTADOS

A lo largo del proceso de búsqueda y recopilación de información, se llevó a cabo una revisión exhaustiva de la literatura existente sobre la aplicación de técnicas de Machine Learning (ML) en la seguridad de la información en el sector educativo. Para minimizar la posibilidad de omitir artículos relevantes, se realizaron búsquedas en bases de datos académicas reconocidas, como IEEE Xplorer, utilizando términos específicos como "Machine Learning", "seguridad de la información", "vulnerabilidades", "sector educativo" y "detección de intrusiones".Los resultados de las búsquedas se organizaron en una hoja de cálculo de Microsoft Excel, donde se registraron los artículos encontrados, se evaluaron y se resolvieron discrepancias durante el proceso de revisión. El diagrama de flujo PRISMA muestra el proceso detallado de búsqueda, recopilación y selección de datos.

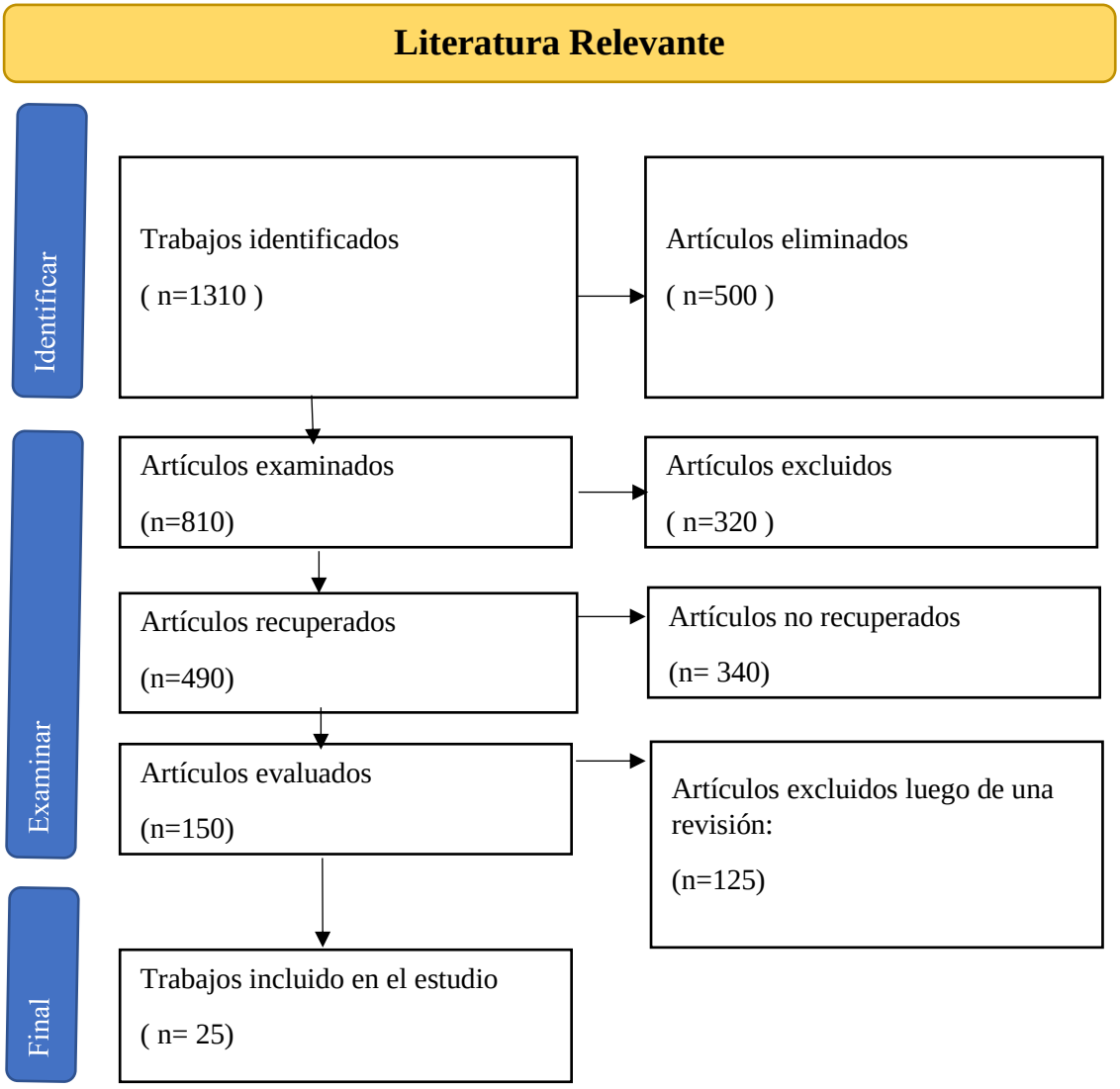
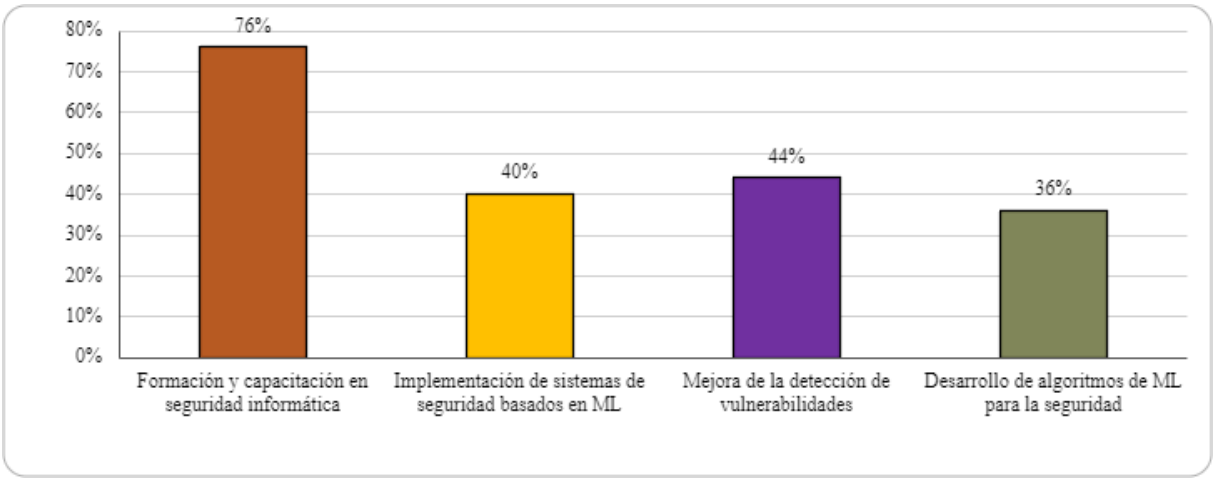


Figura 3.

Distribución de temática en los estudios revisados



Formación y capacitación en seguridad informática

Un 76 % de los estudios analizados destaca la importancia de la información y capacitación en seguridad informática para mejorar la protección de los datos en las instituciones educativas (ver Fig. 1). Estos estudios indican que la capacitación con del personal y los estudiantes en técnicas de seguridad es crucial para prevenir incidentes de seguridad y garantizar una repuesta efectiva ante los posibles ataques cibernéticos (Abraham & Bindu, 2021).

Implementación de sistemas de seguridad basados en ML

Un 40% de los artículos revisados abordan la implementación de sistemas de seguridad que utilizan Machine Learning para detectar y prevenir ataques cibernéticos (ver Fig. 1). Estos sistemas pueden analizar grandes volúmenes de datos en tiempo real, identificar patrones anómalos y predecir amenazas potencia, lo que mejora mucho la capacidad de repuesta ante incidentes de seguridad (Gupta et al., 2021).

Mejora de la detección de vulnerabilidades

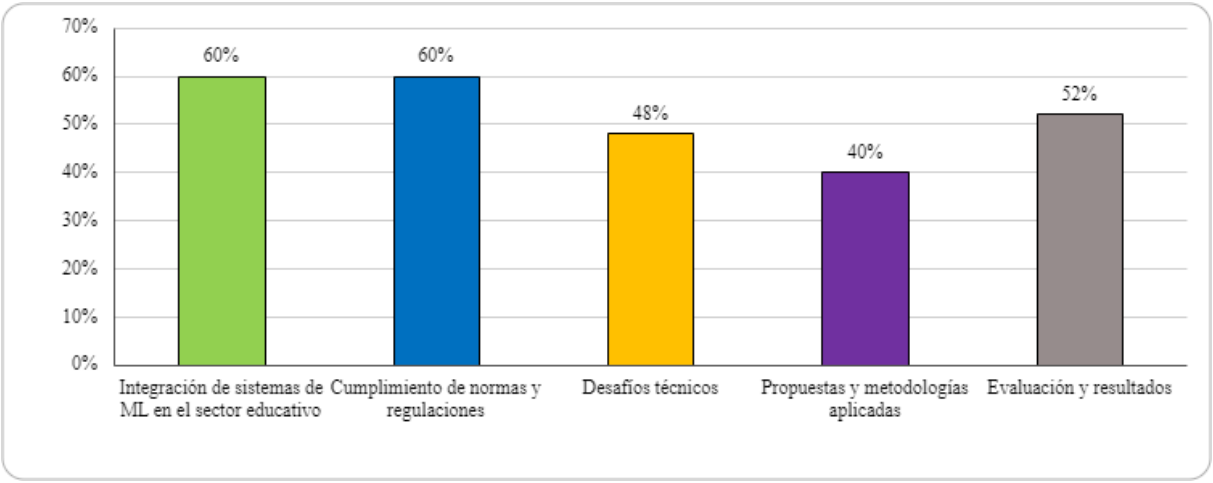
El 44% de los estudios se centran en la mejora de la detección de vulnerabilidades a través de algo avanzados de Machine Learning (ver Fig. 1). La capacidad de estos algoritmos para identificar vulnerabilidades de manera proactiva es esencial para proteger los sistemas educativos contra ataques cibernéticos cada vez más sofisticados (Krishna et al., 2023b).

Desarrollo de algoritmos de ML para la seguridad

Un 36% de los artículos analizados investigan el desarrollo de nuevos algoritmos de Machine Learning específicamente diseñados para aplicaciones de seguridad (ver Fig. 1). Estos desarrollos incluyen técnicas como redes neuronales profundas y aprendizaje automático continuo, que han demostrado ser altamente efectivos en la detección de amenazas complejas (Brintha et al., 2023) .

Figura 2.

Evaluación de los desafíos técnicos y metodológica aplicada



Integración de sistemas de ML en el sector educativo

El 60% de los estudios discuten la integración de sistemas de Machine Learning en el sector educativo, destacando los beneficios y desafíos asociados (ver Fig. 2). La integración efectiva de estas tecnologías puede mejorar significativamente la seguridad de la información, pero también presenta desafíos en términos de gestión de datos y adaptación a nuevas amenazas (Hossain et al., 2023).

Cumplimiento de normas y regulaciones

Un 60% de los estudios revisados enfatizan la importancia de cumplir con normas y regulaciones de seguridad, como ISO/IEC 27001 y 27002 (ver Fig. 2). Cumplir con estos estándares asegura que las instituciones educativas mantengan un alto nivel de seguridad y protección de datos (Kaipu et al., 2023).

Desafíos técnicos y metodología aplicadas

Los estudios también destacan diversos desafíos técnicos, 48% y proponen metodologías aplicadas 40% para mejorar la seguridad en el sector educativo (ver Fig. 2). Estos desafíos incluyen la gestión de grandes volúmenes de datos y la necesidad de desarrollar metodologías robustas para la implementación de sistemas de ML (Y. Liu et al., 2021).

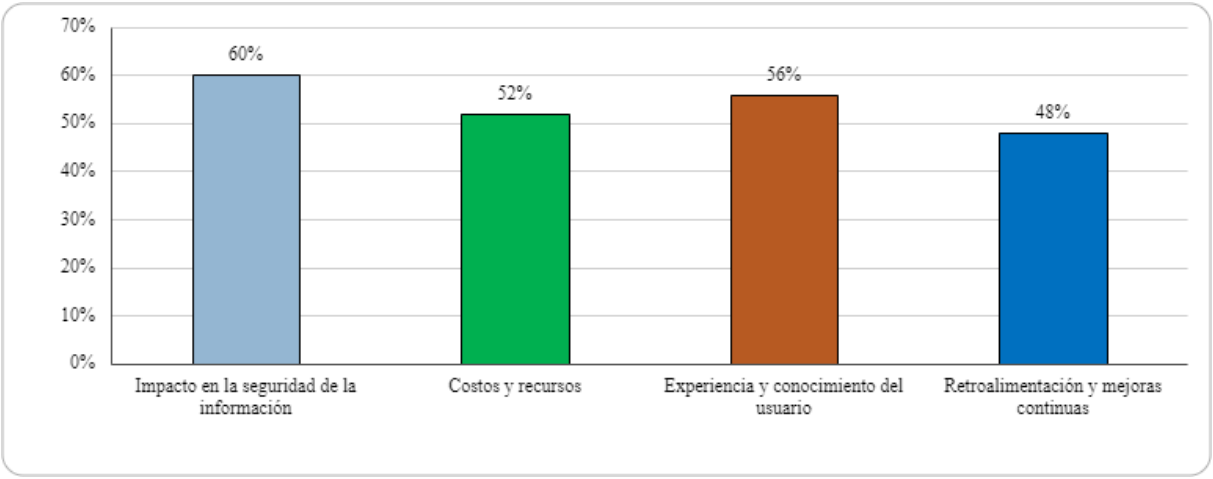
Evaluación y resultados

El 52% de los artículos revisados presentan evaluaciones detalladas y resultados de la aplicación de técnicas de Machine Learning en la seguridad de la información (ver Fig. 2). Estas evaluaciones son

cruciales para entender la efectividad de los sistemas de ML y para identificar áreas de mejora (Lourens et al., 2022).

Figura 3.

Resultado de la implementación de Machine Learning en la seguridad



Impacto en la seguridad de la información

El 60% de los estudios revisados destacan el impacto significativo de las técnicas de Machine Learning en la mejora de la seguridad de la información (ver Fig. 3). Estos estudios muestran cómo las técnicas de ML pueden mejorar la detección y respuesta a incidentes de seguridad ((Aledam & Al-Latteef, 2024).

Costos y recursos

Un 52% de los estudios analizan los costos y recursos necesarios para la implementación efectiva de sistemas de Machine Learning en la seguridad de la información (ver Fig. 3). La inversión en tecnologías de ML puede ser alta, pero los beneficios en términos de seguridad y eficiencia justifican estos costos (Singhal et al., 2023).

Experiencia y conocimiento del usuario

El 56% de los artículos enfatizan la importancia de la experiencia y el conocimiento del usuario en la efectividad de las soluciones de Machine Learning para la seguridad (ver Fig. 3). La formación y capacitación continua del personal es fundamental para maximizar los beneficios de las tecnologías de ML (Wajahat et al., 2024).

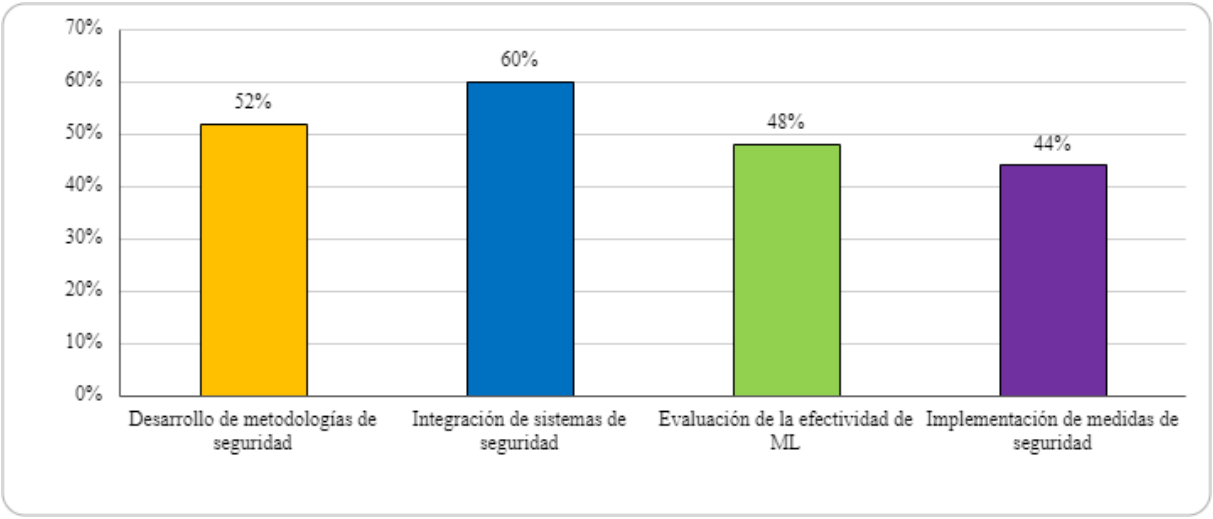
Retroalimentación y mejoras continuas

Un 48% de los estudios revisados mencionan la importancia de la retroalimentación y las mejoras continuas en la implementación de sistemas de Machine Learning (ver Fig. 3). La

retroalimentación regular permite ajustar y mejorar los sistemas de ML de manera continua (Pavithra et al., 2023).

Figura 4.

Metodología y efectividad de ML en la seguridad



Desarrollo de metodología de seguridad

Un 52% de los estudios abordan el desarrollo de metodologías específicas de seguridad utilizando técnicas de Machine Learning (ver Fig. 4). Estas metodologías incluyen el uso de algoritmos avanzados para detectar y prevenir amenazas cibernéticas (Kushal et al., 2024).

Integración de sistemas de seguridad

El 60% de los estudios revisados destacan la importancia de la integración de sistemas de seguridad, enfatizando cómo las técnicas de Machine Learning pueden ser combinadas con otras soluciones de seguridad para ofrecer una protección más robusta y efectiva (ver Fig. 4). La integración de diferentes sistemas de seguridad no solo mejora la detección y respuesta a amenazas, sino que también facilita una gestión centralizada y coherente de la seguridad de la información en las instituciones educativas (Kaipu et al., 2023).

Evaluación de la efectividad de ML

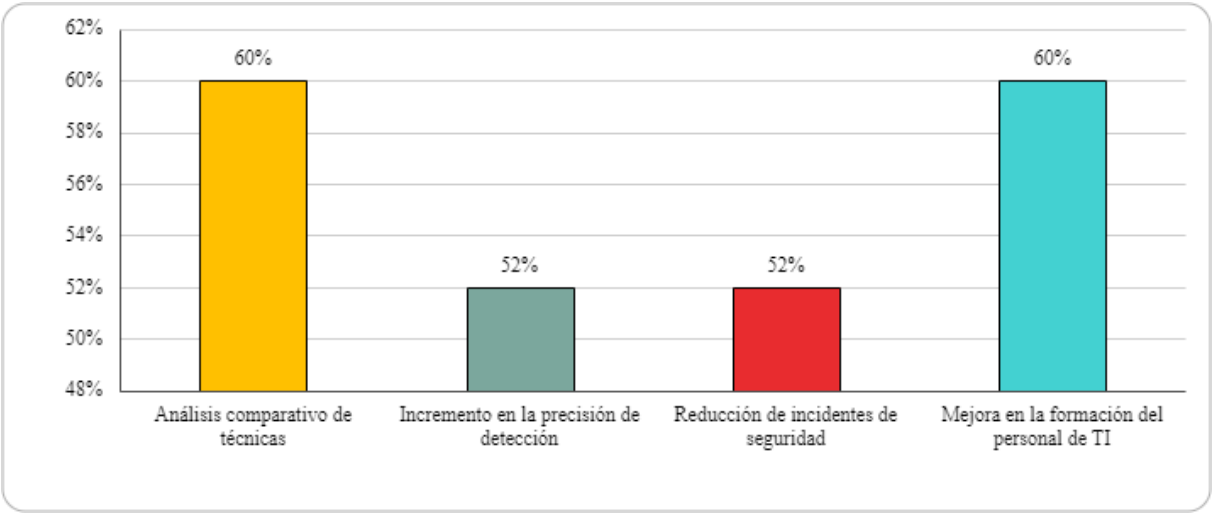
El 48% de los artículos revisados se centran en la evaluación de la efectividad de las técnicas de Machine Learning implementadas (ver Fig. 4). La evaluación constante es esencial para garantizar que las soluciones de ML sigan siendo efectivas ante nuevas amenazas (Rajeshkumar et al., 2022).

Implementación de medida de seguridad

Un 44% de los estudios discuten la implementación de medidas de seguridad basadas en los resultados de las evaluaciones de ML (ver Fig. 4). Estas medidas incluyen la adopción de mejores prácticas y estándares internacionales para mejorar la seguridad de la información (Priya et al., 2022).

Figura 5.

Análisis comparativos y mejoras en la detección de detección



Análisis comparativo de técnicas

El 60% de los artículos incluyen un análisis comparativo de diferentes técnicas de Machine Learning utilizadas para la seguridad de la información (ver Fig. 5). Este análisis es crucial para identificar las técnicas más efectivas y adecuadas para diferentes contextos (Mittal et al., 2023).

Incremento de precisión de detección

Un 52% de los estudios destacan el incremento en la precisión de detección de amenazas cibernéticas gracias al uso de Machine Learning (ver Fig. 5). La precisión mejorada permite una detección más rápida y precisa de incidentes de seguridad (Y. Liu et al., 2021) .

Reducción de incidentes de seguridad

El 52% de los artículos analizan la reducción significativa de incidentes de seguridad tras la implementación de técnicas de Machine Learning (ver Fig. 5). La adopción de estas técnicas ha demostrado reducir notablemente la cantidad de incidentes de seguridad (Yun et al., 2023).

Mejora en la formación del personal de TI

Un 60% de los estudios enfatizan la importancia de mejorar la formación del personal de TI para maximizar los beneficios de las tecnologías de Machine Learning (ver Fig. 5). La

capacitación continua es esencial para garantizar que el personal esté preparado para enfrentar nuevas amenazas (Kristallia et al., 2021).

5. DISCUSIÓN

La revisión de la literatura ha demostrado que el Machine Learning (ML) ofrece numerosas ventajas para la detección y prevención de intrusiones en el sector educativo. Las técnicas avanzadas como las Redes Neuronales Artificiales (ANN) y las Máquinas de Soporte Vectorial (SVM) han mostrado una alta eficacia en la identificación de patrones anómalos y la clasificación de comportamientos maliciosos. Sin embargo, la implementación de estas tecnologías también presenta desafíos, como la necesidad de gestionar grandes volúmenes de datos y la capacidad de los sistemas para adaptarse a nuevas amenazas (Gupta et al., 2021; Priya et al., 2022).

Además, la seguridad de los datos sensibles en entornos educativos es crucial, y el ML ha demostrado ser una herramienta efectiva para proteger estos datos contra accesos no autorizados y ciberataques. La capacidad de modelar y predecir ataques cibernéticos mediante ML permite a las instituciones educativas implementar medidas preventivas más eficaces (Aledam & Al-Latteef, 2024).

Los avances recientes en ML han mejorado significativamente la capacidad de las instituciones educativas para detectar y responder a amenazas cibernéticas. No obstante, persisten desafíos, como la gestión de grandes volúmenes de datos y la adaptación continua a nuevas amenazas ((Zhou et al., 2020). La investigación futura debe centrarse en desarrollar métodos más robustos y confiables que puedan mitigar estas debilidades y mejorar la efectividad de los sistemas de detección de intrusiones (Abraham & Bindu, 2021).

6. CONCLUSIÓN

Se llevó a cabo un análisis exhaustivo de la literatura sobre el uso de Machine Learning para mejorar la seguridad de la información en el sector educativo. Las técnicas de ML, como las Redes Neuronales Artificiales (ANN) y las Máquinas de Soporte Vectorial (SVM), han demostrado ser altamente efectivas en la detección y prevención de intrusiones y en la protección de datos sensibles. Sin embargo, la implementación de estas tecnologías también enfrenta desafíos significativos, incluyendo la gestión de grandes volúmenes de datos y la adaptación continua a nuevas amenazas (Kristallia et al., 2021).

La revisión ha revelado que el ML puede modelar y predecir ataques cibernéticos, permitiendo a las instituciones educativas implementar medidas preventivas más eficaces (Abraham & Bindu, 2021; Aledam & Al-Latteef, 2024; Hossain et al., 2023). Además, la capacidad de estas tecnologías para identificar patrones anómalos y comportamientos maliciosos es esencial para proteger los datos sensibles contra accesos no autorizados (Agrawal et al., 2023; Z. Liu, 2021). Los avances recientes, como el uso de algoritmos de aprendizaje profundo y técnicas de clasificación, han mejorado significativamente la capacidad de detección y respuesta a amenazas cibernéticas (Krishna et al., 2023b; Y. Liu et al., 2021). Sin embargo, persisten desafíos, como la necesidad de una mayor robustez en los sistemas y la continua evolución de las amenazas (Kaipu et al., 2023; Kushal et al., 2024).

El análisis de los estudios revisados también ha destacado la importancia de cumplir con las normas y regulaciones internacionales de seguridad, como ISO/IEC 27001 y 27002, para garantizar un alto nivel de protección de los datos (Mittal et al., 2023; Singhal et al., 2023). Además, se identificaron buenas prácticas y recomendaciones para la implementación de ML en la seguridad cibernética educativa, lo que podría guiar futuros desarrollos en este campo (Li et al., 2023; Rajeshkumar et al., 2022).

La investigación también destaca la importancia de la formación continua y la experiencia del usuario en la efectividad de las soluciones de ML (Wajahat et al., 2024). La capacitación del personal es esencial para maximizar los beneficios de estas tecnologías. Asimismo, la implementación de sistemas de ML debe ser acompañada por una evaluación y mejora constante, adaptándose a las nuevas amenazas y optimizando la gestión de seguridad (Pavithra et al., 2023).

Se concluye que, aunque el uso de ML en la seguridad de la información en el sector educativo es prometedor, es esencial una investigación continua para superar los desafíos existentes y maximizar los beneficios de estas tecnologías (Luo, 2022). La implementación efectiva de ML no solo mejorará la protección de los datos sensibles, sino que también optimizará la gestión de seguridad en las instituciones educativas, proporcionando un entorno más seguro para estudiantes y personal (Brintha et al., 2023; Lan, 2024).

REFERENCIAS

Abraham, J. A., & Bindu, V. R. (2021). Intrusion Detection and Prevention in Networks Using Machine Learning and Deep Learning Approaches: A Review. *2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation, ICAECA 2021*. https://doi.org/10.1109/ICAECA52838.2021.9675595

Agrawal, K., Poli, P. K. R., Gupta, Y., Juneja, G., Bopaiah, A. K., & Arani, D. A. (2023). IoT Attack Detection and Prevention Through Machine Learning System. *2023 International Conference on Communication, Security and Artificial Intelligence, ICCSAI 2023*, 103–106. https://doi.org/10.1109/ICCSAI59793.2023.10421216

Aledam, F. M. M., & Al-Latteef, B. M. A. (2024). Enhanced Malware Detection for Mobile Operating Systems Using Machine Learning and Dynamic Analysis. *International Journal of Safety and Security Engineering*, *14*(2), 513–521. https://doi.org/10.18280/ijsse.140218

Alvarado-Salazar, R., & Llerena-Izquierdo, J. (2022). Revisión de la literatura sobre el uso de Inteligencia Artificial enfocada a la atención de la discapacidad visual. *Revista InGenio*, *5*(1), 10–21. https://doi.org/https://doi.org/10.18779/ingenio.v5i1.472

Brintha, N. C., Abinivesh, S., Sivadasan, A., Balasurya, S., & Babu, S. S. H. (2023). Analysis and Detection of Malware using Machine Learning. *2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, 1329–1334. https://doi.org/10.1109/ICSCDS56580.2023.10105104

Gupta, S. K., Tripathi, M., & Grover, J. (2021). Towards an Effective Intrusion Detection System using Machine Learning techniques: Comprehensive Analysis and Review. *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), ICRITO 2021*. https://doi.org/10.1109/ICRITO51393.2021.9596369

Hossain, M. N., Hassan, Md. M., Monir, R. J., Sayeed, Md. S., Wajiha, S., & Wazid Ullah, S. M. (2023). Cyber Security and People: Human Nature, Psychology, and Training Affect User Awareness, Social Engineering, and Security Professional Education and Preparedness. *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 1–5. https://doi.org/10.1109/ICCCNT56998.2023.10307467

Kaipu, C. N. R., Karthik, G., Umadevi, K. S., Koushik, K. K., Pavan Kumar, T., & Kavitha, S. (2023). An Exploration of Evaluating the Performance of Malware Detection in the Cloud Environment. *2023 International Conference on Computer Communication and Informatics (ICCCI)*, 1–6. https://doi.org/10.1109/ICCCI56745.2023.10128577

Krishna, D. V., Kumar, G. S., Kumar, R. L., Manikanta, K., Vamsidhar, E., & Moulana, M. (2023a). Malware Detection using Machine Learning. *2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, 340–344. https://doi.org/10.1109/ICAAIC56838.2023.10141501

Krishna, D. V., Kumar, G. S., Kumar, R. L., Manikanta, K., Vamsidhar, E., & Moulana, M. (2023b). Malware Detection using Machine Learning. *Proceedings of the 2nd International Conference on Applied Artificial Intelligence and Computing, ICAAIC 2023*, 340–344. https://doi.org/10.1109/ICAAIC56838.2023.10141501

Kristallia, R., Setiawan, H., & Sabiya, S. M. (2021). Design and Development Hands-On Vulnerable Web Application as a Software Security Educational Media. *2021 Sixth International Conference on Informatics and Computing (ICIC)*, 1–6. https://doi.org/10.1109/ICIC54025.2021.9632902

Kushal, S., Shanmugam, B., Sundaram, J., & Thennadil, S. (2024). Self-healing hybrid intrusion detection system: an ensemble machine learning approach. *Discover Artificial Intelligence*, *4*(1), 28. https://doi.org/10.1007/s44163-024-00120-9

Lan, Z. (2024). RETRACTED ARTICLE: Network security threat detection technology based on EPSO-BP algorithm. *EURASIP Journal on Information Security*, *2024*(1), 6. https://doi.org/10.1186/s13635-024-00152-9

Li, Q., Liu, B., & Chen, P. (2023). An overview of cybersecurity based on Network Security Situational Awareness and Machine learning. *2023 8th International Conference on Intelligent Computing and Signal Processing, ICSP 2023*, 279–285. https://doi.org/10.1109/ICSP58490.2023.10248496

Liu, Y., Li, J., Liu, B., Gao, X., & Liu, X. (2021). Malware Identification Method Based on Image Analysis. *2021 11th International Conference on Information Technology in Medicine and Education (ITME)*, 157–161. https://doi.org/10.1109/ITME53901.2021.00041

Liu, Z. (2021). Construction of Computer Mega Data Security Technology Platform Based on Machine Learning. *2021 IEEE 4th International Conference on Information Systems and Computer Aided Education, ICISCAE 2021*, 538–541. https://doi.org/10.1109/ICISCAE52414.2021.9590732

López-Chila, R., Llerena-Izquierdo, J., Sumba-Nacipucha, N., & Cueva-Estrada, J. (2024). Artificial Intelligence in Higher Education: An Analysis of Existing Bibliometrics. *Education Sciences*, *14*(1). https://doi.org/10.3390/educsci14010047

Lourens, M., Dabral, A. P., Gangodkar, D., Rathour, N., Tida, C. N., & Chadha, A. (2022). Integration of AI with the Cybersecurity: A detailed Systematic review with the practical issues and challenges. *2022 5th International Conference on Contemporary Computing and Informatics (IC3I)*, 1290–1295. https://doi.org/10.1109/IC3I56241.2022.10073040

Luo, Y. (2022). Analysis and Research of Network Information Security Evaluation Model Based on Machine Learning Algorithm. *Proceedings of the 4th IEEE Eurasia Conference on IoT, Communication and Engineering 2022, ECICE 2022*, 193–196. https://doi.org/10.1109/ECICE55674.2022.10042828

Mittal, S., Mishra, A. K., Wazid, M., Singh, D. P., Das, A. K., & Shetty, S. (2023). Multiclass Classification Approaches for Intrusion Detection in IoT-Driven Aerial Computing Environment. *GLOBECOM 2023 - 2023 IEEE Global Communications Conference*, 2160–2165. https://doi.org/10.1109/GLOBECOM54140.2023.10436894

Pavithra, B., Vinola, C., Mishra, N., & Naveen, G. (2023). Cloud Security Analysis using Machine Learning Algorithms. *Proceedings of the 2023 2nd International Conference on Augmented Intelligence and Sustainable Systems, ICAISS 2023*, 704–708. https://doi.org/10.1109/ICAISS58487.2023.10250594

Priya, N. S., Meyyappan, S., Balasubramanian, K. N., & Pruthiev, A. S. (2022). Network Attack Detection using Machine Learning. *8th International Conference on Advanced Computing and*

Communication Systems, ICACCS 2022, 342–346. https://doi.org/10.1109/ICACCS54159.2022.9785263

Rajeshkumar, K., Dhanasekaran, S., & Vasudevan, V. (2022). Applications of Machine Learning Algorithms for HDFS Big Data Security. *2022 International Conference on Computer Communication and Informatics, ICCCI 2022*. https://doi.org/10.1109/ICCCI54379.2022.9740908

Sanchez-Romero, J., & Llerena-Izquierdo, J. (2023). Revisión de la literatura sobre el uso del aprendizaje profundo enfocado en sistemas de inspección ópticos automatizados para la detección de defectos superficiales en el sector de la manufactura. *Revista InGenio*, *6*(2), 1–19. https://doi.org/10.18779/ingenio.v6i2.680

Singhal, S., Srivastava, R., Shyam, R., & Mangal, D. (2023). Supervised Machine Learning for Cloud Security. *2023 6th International Conference on Information Systems and Computer Networks, ISCON 2023*. https://doi.org/10.1109/ISCON57294.2023.10112078

Wajahat, A., He, J., Zhu, N., Mahmood, T., Nazir, A., Ullah, F., Qureshi, S., & Dev, S. (2024). Securing Android IoT devices with GuardDroid transparent and lightweight malware detection. *Ain Shams Engineering Journal*, *15*(5), 102642. https://doi.org/10.1016/j.asej.2024.102642

Yun, K., Jin, Y., Huang, Q., & Wang, Q. (2023). A Network Security Approach based on Machine Learning. *2023 IEEE International Conference on Integrated Circuits and Communication Systems, ICICACS 2023*. https://doi.org/10.1109/ICICACS57338.2023.10100204

Zerega-Prado, J., & Llerena-Izquierdo, J. (2022). Arquitectura de consolidación de la información para seguros de la salud mediante Big Data. *Memoria Investigaciones En Ingeniería*, *0*(23 SE-Artículos). https://doi.org/10.36561/ING.23.3

Zhou, H., Yang, X., Pan, H., & Guo, W. (2020). An Android Malware Detection Approach Based on SIMGRU. *IEEE Access*, *8*, 148404–148410. https://doi.org/10.1109/ACCESS.2020.3007571