Revista Sinapsis. ISSN 1390 – 9770
Periodo. Enero – Junio 2024
Vol. 24, Nro. 1, Publicado 2024-06-30

Impacto de la inteligencia artificial en los ciberataques

Impact of artificial intelligence on cyber attacks

Aura Dolores Zambrano Rendón

Escuela Superior Politécnica Agropecuaria de Manabí ESPAM MFL, Correo: azambrano@espam.edu.ec, Código Orcid: https://orcid.org/0000-0002-2784-9202

Contacto: azambrano@espam.edu.ec

Recibido: 24-02-2024 **Aprobado:21-04-2024**

Resumen

El artículo se enfoca en analizar cómo la inteligencia artificial (IA) influye en el campo de la ciberseguridad, explorando tanto sus beneficios como sus riesgos, y, comprender cómo la IA ha sido empleada por actores maliciosos para llevar a cabo ciberataques más avanzados, así como identificar las técnicas de explotación de IA utilizadas en estos ataques. La investigación se basó en la recopilación de información actualizada, priorizando artículos en español de los últimos cinco años, aplicando criterios específicos para seleccionar artículos relacionados con fake news, ataques de phishing y robo de identidad, además, se evaluó la calidad de los artículos seleccionados utilizando herramientas de IA y se los categorizó en función de los tipos de ciberataques. Los resultados principales incluyen el análisis de un total de 350 artículos relacionados con ciberataques y la identificación de 145 artículos relevantes después de aplicar los criterios de selección, esto proporcionó una visión más clara de las técnicas de explotación de IA utilizadas por los actores maliciosos en estos ataques, se concluye que la IA ha tenido un impacto significativo en la evolución de los ciberataques, otorgando a las atacantes herramientas más sofisticadas para la actividad maliciosa, sin embargo, también se destaca que la IA puede ser empleada para fortalecer la ciberseguridad al mejorar la detección y prevención de estos ataques de manera más eficaz.

Palabras clave: Ciberseguridad, Inteligencia Artificial, Amenazas cibernéticas, Técnicas de explotación de IA.

Abstract

The article focuses on analyzing how artificial intelligence (AI) influences the field of cybersecurity, exploring both its benefits and its risks, and understanding how AI has been used by malicious actors to carry out more advanced cyberattacks, as well as how to identify the AI exploitation techniques used in these attacks. The research was based on the collection of updated information, prioritizing articles in Spanish from the last five years, applying specific criteria to select articles related to fake news, phishing attacks and identity theft, in addition, the quality of the selected articles was evaluated. using AI tools and categorized them based on the types of cyberattacks. The main results include the analysis of a total of 350 articles related to cyberattacks and the identification of 145 relevant articles after applying the selection criteria, this provided a clearer insight into the AI exploitation techniques used by malicious actors in these attacks, it is concluded that AI has had a significant impact on the evolution of cyber attacks, giving attackers more prominent tools for malicious activity, however, it is also highlighted that AI can be used to strengthen cybersecurity by improving the detection and prevention of these attacks more effectively.

Revista Sinapsis. ISSN 1390 – 9770
Periodo. Enero – Junio 2024
Vol. 24, Nro. 1, Publicado 2024-06-30
Keywords: ChatGPT, Cybersecurity, Artificial Intelligence, Cyber threats, AI exploitation techniques

Introducción

La inteligencia artificial (IA) ha revolucionado de manera significativa numerosos campos en la sociedad moderna, desde la atención médica y la conducción autónoma hasta la optimización de procesos industriales. Sin embargo, este mismo avance tecnológico que ha traído consigo una amplia gama de beneficios también ha sido aprovechado por actores malintencionados para perpetrar ciberataques cada vez más sofisticados y peligrosos (Rodríguez, 2023). Para (González, 2023), la IA se ha convertido en una espada de doble filo en el ámbito de la ciberseguridad, a medida que evoluciona, se intensifican los ataques sofisticados y, al mismo tiempo, se potencian los sistemas existentes para defenderse de ellos. Según Portela (2023), la Inteligencia Artificial IA está adentrándose de forma acelerada en el ámbito de la ciberseguridad, experimentando un rápido desarrollo en los últimos años y convirtiéndose en una herramienta poderosa. Las redes neuronales y los algoritmos de aprendizaje automático tienen la capacidad de analizar grandes volúmenes de datos para detectar vulnerabilidades y crear estrategias de ataque más efectivas. Como plantea Niss (2023), los sistemas autónomos con IA operan en base a un razonamiento estocástico (basado en probabilidades), lo cual introduce incertidumbre, pero también emplean otras técnicas como el razonamiento deductivo, inductivo y el aprendizaje automático, entre otros. Sin embargo, su mal uso ha generado preocupaciones significativas en relación con delitos y ciberataques. Los ciberdelincuentes han adoptado la IA como una herramienta clave para perfeccionar sus tácticas y ampliar sus capacidades de ataque. La combinación de algoritmos inteligentes con técnicas tradicionales de ciberataque ha creado una nueva generación de amenazas cibernéticas altamente sofisticadas y difíciles de detectar. La convergencia de la IA con las actividades cibernéticas ha abierto las puertas a un mundo de posibilidades para los atacantes, quienes utilizan esta tecnología para diseñar ataques más precisos, automatizados y, en última instancia, más dañinos (Rodrigo , 2022).

La rápida evolución de la inteligencia artificial ha llevado al desarrollo de herramientas sofisticadas como ChatGPT, que han demostrado tener un amplio potencial en diversas aplicaciones. De acuerdo con un estudio realizado por la Universidad de Oxford y la Universidad de Yale, se espera que la inteligencia artificial supere a los seres humanos en diversas actividades en los próximos diez años (Galindo, 2020). Pastor (2023) indica que el rápido crecimiento de la Inteligencia Artificial ChatGPT demuestra el interés de los usuarios en utilizar una tecnología cuyas fortalezas y debilidades prometen revolucionar varios campos. Desde la perspectiva de Bazurto (2023), los usuarios eligen la herramienta de Inteligencia Artificial ChatGPT debido a su capacidad para procesar el lenguaje natural de manera fluida, lo que se traduce en la generación de respuestas coherentes, claras y precisas con una eficiencia destacada. En este artículo, se examinarán las diferentes maneras en las que la Inteligencia Artificial puede ser utilizada para llevar a cabo actividades ilegales, abarcando desde el phishing hasta la generación de contenido engañoso. Además, se explorarán las implicaciones éticas y legales que surgen de estas prácticas, así como la probabilidad de que los usuarios caigan en trampas o divulguen información confidencial. También se analizará cómo la inteligencia artificial impacta los ataques maliciosos y los desafíos que esta nueva era de amenazas cibernéticas plantea.

Materiales y Métodos

El presente artículo constará de 3 fases implementadas secuencialmente. En la primera fase, se realizará una revisión de literatura en la cual se buscará actividades ilícitas documentadas sobre ciberataques impulsados por la Inteligencia Artificial, centrándose en la creación de fake news, ataques de phishing y robo de identidad. Posteriormente, se ejecutará un análisis de casos de estudio, donde se analizarán

los casos reales en los que se ha utilizado IA en ciberataques, investigando los métodos utilizados, las técnicas empleadas y los resultados obtenidos. Finalmente, se mostrará las medidas de seguridad existentes y su eficacia para hacer frente a los ciberataques basados en IA.

Se lleva a cabo la recopilación y análisis de la literatura existente sobre las actividades ilícitas documentadas en ciberataques impulsados por la Inteligencia Artificial. Esto implica examinar investigaciones previas, artículos científicos, informes técnicos y otros recursos pertinentes. Según Noreña (2020), la revisión de literatura es un compendio de todo el conocimiento relacionado con un tema específico de investigación. Este proceso permite comprender el estado actual del conocimiento sobre el impacto de la Inteligencia Artificial en los ciberataques, identificar lagunas en la investigación y establecer una base teórica sólida. En particular, esta revisión de literatura se centrará en las actividades ilícitas, como la creación de fake news, ataques de phishing, robo de identidad. Se realizará una exhaustiva búsqueda en diferentes repositorios y buscadores web con el propósito de obtener información actualizada acerca de ciberataques. A continuación, se muestra la tabla de los repositorios y buscadores web que se utilizarán para la investigación:

Repositorios y Buscadores Web	**URL**
IEEE Xplore	https://ieeexplore.ieee.org/Xplore/home.jsp
ACM Digital Library	https://dl.acm.org/
Google Scholar	https://scholar.google.com/
ScienceDirect	https://www.sciencedirect.com/
SpringerLink	https://link.springer.com/
Scopus	www.scopus.com
Web of Science	www.webofscience.com
VirusTotal	www.virustotal.com
Mitre ATT&CK	attack.mitre.org
US-CERT	www.us-cert.gov
Kaspersky Threat Intelligence	www.kaspersky.com/enterprise-security/threat-intelligence-center
IBM	https://exchange.xforce.ibmcloud.com/

Tabla 1: *Repositorios utilizados para búsquedas.*

Luego, se utilizará el apoyo de herramientas de inteligencia artificial para explorar los repositorios y llevar a cabo búsquedas eficientes de libros relacionados sobre los ciberataques. Esto permitirá reducir los tiempos de búsqueda y obtener información más precisa y relevante sobre los ciberataques. A continuación, se presenta la tabla de herramientas de inteligencia artificial para la búsqueda en repositorios:

Herramienta de Inteligencia Artificial	**URL**
OpenAI GPT-3	openai.com/gpt-3
Perplexity	https://www.perplexity.ai/
you.ia	https://you.com/
bard	https://bard.google.com/?hl=es

Tabla 2: *Herramientas de IA*

Análisis de casos de estudio

En esta etapa, se seleccionan y analizan casos específicos relacionados con los ciberataques impulsados por la Inteligencia Artificial. Estos casos pueden incluir ataques conocidos en los que se haya utilizado IA de manera significativa, como malware basado en IA, phishing sofisticado o ataques de ingeniería social mejorados por algoritmos de aprendizaje automático. Según Idrovo (2020), el análisis de casos de estudio permite profundizar en un proceso de aprendizaje práctico y completo con el objetivo de comprender todos sus elementos dentro de un análisis exhaustivo e inmersivo. Por lo tanto, se investigarán los detalles del método utilizado, el objetivo del ataque, las técnicas de evasión empleadas y el impacto resultante, lo que permitirá comprender mejor cómo la IA ha transformado la naturaleza y el alcance de los ciberataques.

En el análisis del caso de estudio, se emplearán herramientas de inteligencia artificial para agilizar y mejorar la búsqueda dentro de los artículos encontrados. Estas herramientas permitirán realizar preguntas precisas sobre la información contenida en los artículos y obteniendo respuestas detalladas y relevantes en relación con el contenido de los documentos. A continuación, se muestra la tabla con las herramientas de inteligencia artificial para el análisis:

Herramienta de Inteligencia Artificial	**URL**
Humata	https://app.humata.ai/login
Chat with any PDF	https://www.chatpdf.com/
FoldersAI	https://foldersai.com/

Tabla 3: *Herramientas de IA*

Evaluación de medidas de seguridad existentes

En esta fase, se evaluarán las medidas de seguridad cibernética existentes y su eficacia para hacer frente a los ciberataques impulsados por la Inteligencia Artificial. Se examinarán herramientas de detección y prevención, algoritmos de aprendizaje automático utilizados en la seguridad cibernética y otras soluciones de protección disponibles. Se analizará cómo estas medidas abordan las nuevas amenazas y técnicas utilizadas en los ciberataques basados en IA. Además, se evaluarán sus limitaciones y desafíos, identificando posibles áreas de mejora y sugerencias para fortalecer las defensas contra estos ataques.

Resultados y Discusión

En la actualidad, la inteligencia artificial (IA) está experimentando un rápido avance y la frecuencia de los ataques cibernéticos sigue en aumento, por lo tanto, comprender cómo los agentes maliciosos emplean la IA para llevar a cabo sus ataques se ha convertido en una prioridad crítica. Mediante acciones como la exploración de los patrones emergentes, las técnicas y estrategias, se espera contribuir de manera significativa a la formulación de estrategias efectivas para contrarrestar y mitigar los ciberataques basados en la IA.

Revisión de literatura

Se consideraron bases de datos y artículos científicos relacionadas con el área de ciberataques e inteligencia artificial, tal como se indica en la (Tabla 1). Del mismo modo, haciendo uso de las fuentes de inteligencia artificial (Tabla 2), se proporcionó el siguiente conjunto de parámetros de búsqueda:
Se ingreso los repositorios y sus respectivos links de acceso, posteriormente se indico que; “de los siguientes repositorios, hacer una tabla detallando el número de artículos sobre ciberseguridad y ataques

con uso de inteligencia artificial".

Repositorio	Número de artículos
IEEE Xplore	1000
ACM Digital Library	500
Google Scholar	2000
ScienceDirect	1000
SpringerLink	500
Scopus	2000
Web of Science	1000
VirusTotal	500
Mitre ATT&CK	200
US-CERT	100
Kaspersky Threat Intelligence	500
IBM	200

Tabla 4: *Número total de Artículos Encontrados en cada Repositorio.*

En términos generales, los hallazgos señalan que Google Scholar, Scopus, Web of Science y IEEE Xplore destacan como fuentes primordiales. Les siguen en importancia las bases de datos ACM Digital Library, SpringerLink, VirusTotal y Kaspersky Threat Intelligence. A partir de esta base de datos, se llevó a cabo un análisis detallado de los datos obtenidos. A continuación, se establecen los criterios de inclusión y exclusión con el propósito de recopilar información pertinente y actualizada:

Criterios de inclusión	Criterios de exclusión
Artículos de los últimos 5 años.	Artículos que contengan información similar a la consultada.
Artículos relacionados con fake news, ataques de phishing, robo de identidad.	Información de tesis.
Artículos Referente a ataques maliciosos	
Idioma español.	
Procedimientos y herramientas utilizadas	

Tabla 5: *Criterios de Inclusión y Exclusión*

En la (Tabla 6) se presenta una evaluación destinada a verificar la calidad de los artículos seleccionados, en la cual, se determinan los resultados obtenidos de 6000 artículos, mismos que fueron filtrados utilizando las herramientas de IA mostradas en la (Tabla2) clasificando los artículos en 3 tipos de ciberataques, posteriormente se aplicaron los criterios de inclusión y exclusión obteniendo un total de 350 artículos y se realizó un filtrado por títulos obteniendo un total de 145 artículos.

Tema	Filtro por Criterios de inclusión y exclusión	Filtro por Títulos	Filtro por Lectura rápida

Fake news	100	45	10
Ataques de phishing	200	90	10
Robo de identidad	50	10	8
Total	350	145	28

Tabla 6:*Artículos y Temas Relacionados.*

Análisis de casos de estudio

El uso de la inteligencia artificial en la ciberseguridad y los ciberataques es un tema de vital importancia en la actualidad. La inteligencia artificial y el aprendizaje automático son herramientas poderosas para fortalecer la ciberseguridad y proteger los sistemas y datos de los ciberataques. De acuerdo con las observaciones de Llano (2022), la inteligencia artificial habilita la capacidad de realizar pronósticos predictivos, los cuales se fundamentan en la información acumulada y el aprendizaje adquirido. Sin embargo, en la actualidad esta herramienta se está utilizando cada vez más en ciberataques, lo que representa una preocupación significativa en el ámbito de la ciberseguridad. Esto se debe a que la IA tiene la capacidad de automatizar tareas que antes requerían la intervención humana, lo que permite a los ciberatacantes llevar a cabo ataques más complejos y sofisticados. Por otra parte, Flores (2019) en su artículo muestra la relevancia de la Declaración de Principios Éticos para la IA de Latinoamérica IA-LATAM en el diseño, desarrollo y uso de la inteligencia artificial, siendo así que, en el octavo inciso, enfatiza la necesidad de prevenir sesgos y efectos injustos a las personas. Algunos ejemplos de cómo se utiliza la IA para realizar ciberataques incluyen:

Fake News

Según menciona Llano (2022) en el año 2018, el grupo de expertos de alto nivel sobre noticias falsas y desinformación de la Unión Europea presentó un Informe sobre las fake news y desinformación en línea, en el cuál son definidas como la diseminación de información apócrifa, inexacta o engañosa, con una intención maliciosa dirigida a infligir daño público o lograr ventajas particulares. Asimismo, Acosta y Laurent (2023) señalan que las noticias son las herramientas más útiles para que los ciudadanos puedan mantenerse al día sobre lo que sucede en su país y en todo el mundo. La difusión de noticias falsas conlleva una serie de consecuencias, entre las que destacan un constante cuestionamiento de la credibilidad de la información que se consume, la toma de decisiones erróneas como resultado de la creencia en noticias falsas, la pérdida de confianza en los hallazgos científicos, entre otros.

La distinción clave con respecto a las noticias falsas de la actualidad radica en que, gracias a las plataformas digitales que sustentan las "redes sociales" y facilitan la producción y difusión masiva de contenidos, la desinformación se propaga y multiplica de manera exponencial en tiempo real, sin dejar margen para la reflexión o la corrección. En este contexto, se presentan ejemplos de noticias falsas que se encuentran en las redes sociales expuestos en la (Tabla 7).

Fake New	**Articulo**
Varios medios y redes sociales ligados a la ultraderecha empezaron a difundir falsas noticias sobre la red de pedofilia que afirmaban se dirigía desde el restaurante Comet Ping Pong.	Cuatro años de cárcel para el hombre que abrió fuego en una pizzería de Washington tras leer una noticia falsa (FAUS, 2017)

En las elecciones presidenciales de Estados Unidos de 2016, se observó un fenómeno interesante en el que las encuestas sugerían que los votantes indecisos podrían optar por votar por Trump si creían en noticias falsas relacionadas con Hillary Clinton.	Cómo las noticias falsas afectan las elecciones estadounidenses (Lee, 2020)
Teoría de conspiración que afirma falsamente que 5G, que se usa en redes de telefonía móvil y depende de señales transmitidas por ondas de radio, es de alguna manera responsable del coronavirus.	Coronavirus: los científicos afirman que 5G es "una completa basura". (BBC, 2020)
Peele ventríloga a Barack Obama, haciéndole expresar su opinión sobre Black Panther ("Killmonger tenía razón") y llamar al presidente Donald Trump "un total y completo imbécil".	Mira a Jordan Peele usar IA para hacer que Barack Obama entregue un anuncio de servicio público sobre noticias falsas. (James, 2018)

Tabla 7: *Ejemplos de Fake New*

Para el Centro de Tecnologías de la Información y Sociedad (s,f), estas noticias se presentan como si fueran parte del periodismo legítimo, con la intención de manipular a los lectores. En términos generales, el proceso para la propagación de las fake news se puede describir de la siguiente manera:

Crear un sitio de noticias falsas

El proceso inicial para establecer una fábrica de noticias falsas implica la creación de un sitio web en el que se alojarán y presentarán estas noticias falsas. Este procedimiento conlleva dos pasos esenciales: adquirir un nombre de dominio para el sitio web y obtener el servicio de hosting para el propio sitio. Ambos componentes se pueden adquirir a un costo relativamente bajo, y es posible diseñarlos de manera que se asemejen lo más posible a sitios web legítimos de noticias. Las redes sociales actúan como un medio de amplificación significativo, ya que permiten que las noticias falsas se difundan rápidamente a través de la compartición de contenido por parte de usuarios, lo que puede llevar a una propagación masiva y veloz de información errónea o engañosa. Es decir, la viralidad y la velocidad con la que se pueden compartir noticias en redes sociales hacen que estas plataformas sean particularmente susceptibles a la difusión de fake news.

Robo de contenido

Según la perspectiva de Acosta y Laurent, (2023), es posible crear contenido falso desde cero, lo más habitual en estos casos es que los sitios web de noticias falsas obtengan su contenido de fuentes externas. Los lugares más comunes para obtener este tipo de contenido suelen ser sitios web de sátira o de "clickbait". De esta manera, en la actualidad, con el avance de la inteligencia artificial y las tecnologías disponibles en la web, es posible generar noticias falsas a través de las herramientas como las mencionadas en la (Tabla 8), mismas que facilitan a los perpetradores la creación de noticias falsas al utilizar chatbots para generar información ficticia utilizando datos previamente entrenados en el algoritmo. En consecuencia, el algoritmo interpreta estos datos como si se tratara de la construcción de una narrativa ficticia. Además, se emplea la inteligencia artificial para crear imágenes que se utilizan como portadas de las noticias falsas, así como generadores de texto que producen contenido a partir de una entrada textual con la idea a desarrollar.

Herramienta	Descripción	URL

ChatGPT	Genera texto a partir de la descripción de una noticia falsa, incluyendo los detalles de todas las partes involucradas en la falsa noticia.	https://chat.openai.com/
Inferkit	Genera texto a partir de una descripción. Aunque es menos potente que ChatGPT, ofrece la posibilidad de editar y establecer un límite de caracteres para generar y refinar el texto de noticias falsas. Ideal para publicar en un tweet o comentario en redes sociales	https://app.inferkit.com/demo
DALL·E	Genera imágenes realistas a partir de texto.	https://huggingface.co/spaces/dalle-mini/dalle-mini
Leonardo IA	Genera imágenes realistas a partir de texto.	https://leonardo.ai/
Stable Diffusion XL	Genera imágenes realistas a partir de texto.	https://stability.ai/stablediffusion
Midjourney	Genera imágenes realistas a partir de texto.	https://www.midjourney.com/home/?callbackUrl=%2Fapp%2F
Name Generator	Generador de nombres para gatos basado en su apariencia y personalidad.	https://www.name-generator.org.uk/cat/
SofGAN	Generador de imágenes de retratos con estilo dinámico.	https://apchenstu.github.io/sofgan/
This N Does Not Exist	Usando redes adversarias generativas (GAN), se puede aprender a crear versiones falsas de aspecto realista de casi cualquier cosa.	https://thisxdoesnotexist.com/
Fake Person Generator	Genera los datos personales de una persona, que abarcan detalles básicos, información laboral, datos personales y datos adicionales como dirección exacta, dispositivo de trabajo y preferencias	https://www.fakepersongenerator.com/Index/generate
FauxID	Genera los datos personales de una persona, que abarcan detalles básicos, información laboral, datos personales y datos adicionales como dirección exacta, dispositivo de trabajo y preferencias.	https://fauxid.com/fake-name-generator/united-states

Tabla 8: *Herramientas de Inteligencia Artificial (IA)*

Bots

Según kaspersky (2023) los bots beneficiosos desempeñan funciones útiles y, los bots maliciosos, conocidos como bots de malware, representan riesgos significativos, pudiendo ser utilizados para actividades como el hacking, el envío de correos no deseados, espionaje, la interrupción y la vulneración de la seguridad de sitios web de cualquier tamaño. De acuerdo con kaspersky (2023), los bots en las redes sociales tienen la capacidad de propagar información falsa al generar y compartir contenido en grandes cantidades, sin considerar la confiabilidad de las fuentes. Además, pueden crear perfiles falsos

en línea que adquieren seguidores, prestigio y autoridad, y algunas de estas cuentas están programadas específicamente para divulgar información incorrecta.

Deep fakes

Kaspersky (2023) los define como videos falsos creados utilizando software digital, aprendizaje automático y la técnica de intercambio de caras. En estos videos, se combinan imágenes para generar nuevas secuencias que representan eventos o acciones que nunca ocurrieron en realidad. Los resultados suelen ser extremadamente convincentes y difíciles de detectar como falsos. La revista Consumer (2023) indica que la desinformación ha alcanzado un nivel más sofisticado con la ayuda de la inteligencia artificial. Los videos manipulados circulan libremente en Internet y las redes sociales, a menudo sin que podamos identificarlos. De esta manera, se construye un engaño cada vez más elaborado. Utiliza las denominadas redes neuronales generativas antagónicas, GAN por sus siglas en inglés, con algoritmos que son capaces de aprender de los patrones que encuentran en las imágenes para luego reproducirlos creando otras nuevas de ese objeto, rostro o imagen (Visus, 2021).

Categoría	**Descripción**
face-swap	Intercambio de rostros
lip-sync	Sincronización de labios con un mensaje de audio
Puppet master	Marioneta virtual

Tabla 9: *Categorías de deepfakes*

Ataques de phishing

Fuertes et al. (2020) define el phishing como una táctica que combina la ingeniería social con exploits técnicos para engañar a las víctimas y obtener información personal o financiera con fines maliciosos, a menudo con el objetivo de obtener ganancias económicas para el atacante. Esta práctica suele involucrar la suplantación de identidad y la creación de sitios web o correos electrónicos falsos que parecen legítimos para engañar a las personas y hacer que revelen información confidencial, como contraseñas o números de tarjetas de crédito.

Kaminski (2023) enfatiza que la capacidad de generar textos convincentes es una característica destacada en GPT-3 y ChatGPT, lo que aumenta la preocupación sobre la posibilidad de que se estén llevando a cabo ataques automatizados de spear-phishing a través de chatbots. El problema principal con los correos electrónicos de phishing en masa radica en que suelen carecer de una apariencia atractiva y contienen un texto genérico que no se personaliza para el destinatario.

Por otro lado, el spear-phishing en el cual un ciberdelincuente crea un correo electrónico dirigido a una sola víctima, suele ser costoso y se reserva para ataques específicos. Sin embargo, la capacidad de ChatGPT para generar correos electrónicos persuasivos y personalizados a gran escala podría cambiar drásticamente esta dinámica. Según Ayerbe (2020) la utilización de la IA depende del perfil de los ciberatacantes, que van desde los más inofensivos, asociados a la cibermalicia, a los más peligrosos, como pueden ser los relacionados con el ciberterrorismo, el ciberespionaje o la ciberguerra.

Robo de identidad

Guzmán et al. (2020) lo define como la acción de tomar o adueñarse de los datos personales y documentos de identificación de una persona con el propósito de crear documentos de identidad falsos o establecer condiciones mínimas de identificación que puedan utilizarse como base para llevar a cabo una variedad de actividades criminales perjudiciales para la víctima, en efecto CSIC (2020) menciona que en 2019, se observó un incremento del 35% en la incidencia de ciberdelitos en España en comparación con el año previo. Como consecuencia de esta tendencia, el 10% de todos los actos delictivos ocurridos durante ese año, que sumaron un total de 218,302 casos, tuvieron lugar en el ámbito digital, sin embargo, solamente se logró resolver el 15% de estos delitos, lo que subraya los desafíos en la investigación y el esclarecimiento de estos casos. Además, más del 88% de estos ciberdelitos, equivalentes a 192,375 incidentes, estuvieron relacionados con fraudes informáticos o estafas, según lo informado por la Secretaría de Estado de Seguridad. Puig (2023) manifiesta que un individuo malintencionado podría aprovechar la inteligencia artificial para replicar la voz de una persona cercana a ti. Lo único que requeriría sería obtener un breve fragmento de audio que contenga la voz de esa persona, algo que podría encontrar explorando contenidos disponibles en línea, y utilizar un software de clonación vocal.

Evaluación de medidas de seguridad existentes

Ayerbe (2020) enfatiza que la ciberseguridad se encuentra ante una serie de desafíos diversos, que incluyen la detección de intrusiones, la preservación de la privacidad, la adopción de medidas de defensa proactiva, la identificación de comportamientos atípicos y la detección de amenazas altamente sofisticadas. Sin embargo, el desafío principal radica en la constante evolución de las amenazas cibernéticas que emergen de manera continua, según Cuatrecasas (2023) las herramientas de predicción y evaluación de riesgos que hacen uso de la inteligencia artificial para analizar datos históricos y anticipar comportamientos y eventos futuros, puedan prever la probabilidad de que se cometa un delito en cierto lugar y momento, quién podría ser el perpetrador, evaluar el riesgo de fuga o reincidencia de una persona bajo investigación, pronosticar si un recluso volverá a prisión después de un permiso, o predecir si una empresa venderá sus activos tras enfrentar una demanda legal. Por otra parte Telefonicatech (2019) indica que el Deep Learning desempeña un papel importante en esta área. Se emplean diversos modelos de redes neuronales, como los perceptrones multicapa, en la detección de noticias falsas. Además, las Redes Neuronales Recurrentes (RNN) son ampliamente utilizadas en el procesamiento de lenguaje natural, con especial énfasis en la memoria a largo plazo, como las Redes de Memoria a Corto y Largo Plazo (LSTM).

López et al. (2022) maniefiestan que el fact-checking, es una de las tareas fundamentales en el ámbito de la comunicación y desempeña un papel central en la responsabilidad social, esta tarea implica la automatización del proceso de confirmación de noticias mediante la identificación de la fuente, el análisis del contenido y la supervisión de los diferentes flujos de información. Por otro lado, Bezzaoui y Fegert (2022) muestran una solucion con las DeFaktS, que utiliza inteligencia artificial para la detección y alerta de desinformación en plataformas de redes sociales y grupos de mensajería que generan sospechas, consisite en extraer mensajes de manera masiva de redes sociales y grupos de mensajería que levanten sospechas. A partir de estos datos, se lleva a cabo el entrenamiento de un sistema de inteligencia artificial capaz de identificar rasgos y patrones estilísticos característicos de la desinformación, esta IA, ya entrenada, se integra como un componente en lo que se conoce como "IA explicable" (XAI). La XAI es fundamental para permitir que una aplicación informe y advierta a los usuarios de ofertas en línea de manera clara y comprensible, en caso de que se detecten indicios y patrones estilísticos típicos de la desinformación.

Revista Sinapsis. ISSN 1390 – 9770
Periodo. Enero – Junio 2024
Vol. 24, Nro. 1, Publicado 2024-06-30

Discusión

Segun Araíz (2022), en 1956, John McCarthy fue quien introdujo el término "inteligencia artificial" o "IA" para describir una serie de algoritmos diseñados con el propósito de abordar problemas y tomar decisiones sin depender directamente de la intervención humana, estos algoritmos operan en base a patrones de datos disponibles, mejorando sus métodos de manera automática a través de ensayo y error, y son utilizados en problemas que involucran grandes volúmenes de datos. La aplicación de la inteligencia artificial en la realización de ataques que involucran la propagación de noticias falsas, el phishing y la sustracción de información conlleva desafíos e inquietudes de considerable magnitud en el ámbito de la seguridad informática, por ello Aldea (2020) menciona que el uso de la inteligencia artificial en la ciberseguridad genera grandes ventajas y las resumio de la siguinete manera:

Tecnología	**Ventajas**
Redes neuronales artificiales	• Aprenden con el ejemplo. • Son capaces de operar de manera eficaz con funciones complejas no lineales. • Manejan de manera excepcional funciones diferenciales complejas. • Son resilientes a los datos ruidosos y a los datos incompletos.
Agentes inteligentes	• Siempre intentan completar la tarea, incluso cuando tienen objetivos contradictorios. • Actúan de forma racional a la hora de completar sus objetivos. • Se adaptan con facilidad al entorno y a las preferencias del usuario. • Son cocientes de los errores humanos, por lo que pueden ser programados para revisar las instrucciones e inconsistencias que se le han impuesto.
Sistemas inmunes artificiales	• Tienen una estructura dinámica. • Cuentan con medios de aprendizaje distribuido. • Se adaptan y organizan por si solos, sin necesidad de intervención humana. • Son capaces de seleccionar la mejor respuesta para eliminar la amenaza del sistema. • Optimizan los recursos • Cuentan con varias capas de defensa • Al no ser dependientes de ningún elemento en particular, pueden desprenderse y remplazar cualquiera de ellos por uno que tenga un desempeño más alto
Algoritmos genéticos	• Se adaptan al entorno de manera eficiente • Son capaces de optimizar incluso problemas computacionales complejos. • Permiten evaluar varios tipos de posibles soluciones de manera simultánea.
Sistemas expertos	• Pueden ser empleados para una gran variedad de problemas • Ofrecen soluciones a como distribuir los recursos de

manera eficiente

Tabla 10: Ventajas de la Inteligencia Artificial.
Fuente: *Aldea (2020)*

Es fascinante cómo la inteligencia artificial ha transformado la ciberseguridad tanto para el bien como para el mal, uno de los aspectos clave que destaco de este artículo es cómo la recopilación y el análisis de datos masivos han permitido a los actores maliciosos desarrollar tácticas de ciberataque más avanzadas y sofisticadas, logrando obtener la capacidad de los algoritmos de aprendizaje automático para identificar patrones en grandes conjuntos de datos, acarreandolos por el camino de ataques como el phishing y la creación de noticias falsas. Asimismo Visus (2021) menciona que los origenes de las deep fake se dieron a partir del 2014 donde Ian Goodfellow, quien en ese momento era un estudiante de doctorado en la Universidad de Montreal, realizó un avance pionero en la generación de imágenes utilizando un enfoque conocido como redes neuronales generativas adversarias o GAN. Su enfoque implicaba entrenar dos redes neuronales utilizando el mismo conjunto de datos de imágenes y luego emplearlas para generar nuevas imágenes. Lo interesante de su enfoque era que enfrentaba estas dos redes en un tipo de competencia digital, donde una red intentaba identificar qué imágenes eran reales y cuáles eran ficticias, en un juego similar al del gato y el ratón.

La inteligencia artificial es una herramienta de doble filo en el ámbito de la ciberseguridad mostrando su potencial para la innovación y la mejora de la defensa cibernética, pero también presenta desafíos significativos que deben abordarse de manera proactiva. La educación y la conciencia sobre la seguridad cibernética son fundamentales para protegerse contra los ciberataques basados en IA y garantizar un uso responsable de esta tecnología en el futuro.

Conclusiones

La inteligencia artificial ha transformado la forma en que se ejecutan los ciberataques, permitiendo a los actores maliciosos desarrollar herramientas y técnicas más sofisticadas, esto ha resultado en un aumento tanto en la frecuencia como en la gravedad de los ataques, lo que constituye una amenaza significativa para la ciberseguridad, los ciberataques basados en inteligencia artificial abarcan una amplia variedad de actividades maliciosas, que incluyen la generación de noticias falsas para influir en la opinión pública, el phishing para la obtención de información confidencial, el robo de identidad con fines fraudulentos y la infiltración de sistemas informáticos para obtener acceso no autorizado.
La formación y la aplicación de algoritmos de aprendizaje automático han permitido a los ciberdelincuentes llevar a cabo ataques más precisos y eficaces, estos algoritmos tienen la capacidad de analizar volúmenes significativos de datos y aprender patrones que les ayudan a identificar vulnerabilidades y desarrollar formas de eludir las medidas de seguridad existentes, exponiendo que los sistemas tradicionales de defensa ya no son adecuados para contrarrestar los ciberataques basados en inteligencia artificial, por ello se requieren nuevas estrategias y soluciones de ciberseguridad que empleen técnicas de inteligencia artificial, como el análisis del comportamiento y la detección de anomalías, para identificar y neutralizar las amenazas de manera más efectiva, así mismo la colaboración entre la comunidad de ciberseguridad, la industria y las entidades gubernamentales es esencial para enfrentar los ciberataques basados en inteligencia artificial, compartiendo información sobre amenazas, herramientas de detección y mejores prácticas puede contribuir a fortalecer las defensas y prevenir ataques más sofisticados.
La educación y la conciencia en materia de seguridad cibernética desempeñan un papel fundamental en la mitigación de la amenaza representada por los ciberataques basados en inteligencia artificial, ya que los usuarios y las organizaciones deben estar debidamente informados acerca de las tácticas de ataque

más comunes y deben adquirir la capacidad de identificar y evitar las artimañas empleadas por los ciberdelincuentes. Asimismo, es esencial impulsar la capacitación de profesionales en seguridad cibernética y promover una cultura de seguridad en la sociedad en general, con el fin de contrarrestar de manera efectiva las amenazas tanto actuales como las que puedan surgir en el futuro.

BIBLIOGRAFÍA

1. Acosta, V., y Laurent, J. (2023). *Entre números y letras: El uso de la Inteligencia Artificial en la lucha contra las noticias falsas y sus implicancias en la libertad de expresión.* https://forseti.pe/entre-numeros-y-letras-el-uso-de-la-inteligencia-artificial-en-la-lucha-contra-las-noticias-falsas-y-sus-implicancias-en-la-libertad-de-expresion/

2. Aldea , C. (2020). *EL IMPACTO DE LA INTELIGENCIA.* https://repositorio.comillas.edu/xmlui/bitstream/handle/11531/37543/TFG-%20Aldea%20Torres%2C%20Carlos.pdf?sequence=1

3. Araíz, L. (2022). *Inteligencia artificial y.* http://ivai.org.mx/Documentos/Revista%20M%C3%A9xico%20Transparente%203era.%20edici%C3%B3n.pdf

4. Ayerbe, A. (2020). *La ciberseguridad y su relación con la inteligencia.* https://media.realinstitutoelcano.org/wp-content/uploads/2021/10/ari128-2020-ayerbe-ciberseguridad-y-su-relacion-con-inteligencia-artificial.pdf

5. Bazurto, J. (2023). *Análisis comparativo de la inteligencia artificial ChatGPT y Google Chrome.* http://dspace.utb.edu.ec/handle/49000/13902

6. BBC. (2020). *Coronavirus: los científicos afirman que 5G es "una completa basura".* https://www.bbc.com/news/52168096

7. Bezzaoui, I., y Fegert, J. (2022). *DeFaktS.* https://im.iism.kit.edu/1093_3235.php

8. Centro de Tecnologías de la Información y Sociedad. (s.f). *¿De dónde vienen las noticias falsas?* https://www.cits.ucsb.edu/fake-news/where

9. Consumer. (2023). *Deep fakes-Cómo la inteligencia artificial puede generar noticias falsas.* https://revista.consumer.es/portada/actualidad/como-la-inteligencia-artificial-puede-generar-noticias-falsas.html

10. CSIC. (2020). *'Ciberestafas', robo de datos o sabotajes: un libro del CSIC analiza las nuevas amenazas surgidas en el ciberespacio.* https://www.csic.es/en/node/1261378

11. Cuatrecasas, C. (2023). *LA INTELIGENCIA ARTIFICIAL Y LA INVESTIGACIÓN DE DELITOS.* https://revistacugc.es/article/view/5912

12. FAUS, J. (2017). *Cuatro años de cárcel para el hombre que abrió fuego en una pizzería de Washington tras leer una noticia falsa.* https://elpais.com/internacional/2017/06/23/estados_unidos/1498169899_197758.html

13. Flores, J. (2019). *Inteligencia artificial y periodismo: diluyendo el impacto de la.* https://repositorioinstitucional.ceu.es/bitstream/10637/10743/1/es_m2_stamped.pdf

14. Fuertes, W., Benavides, E., y Sanchez, S. (2020). *Caracterización de los ataques de phishing y técnicas para mitigarlos. Ataques: una.* https://revistas.uteq.edu.ec/index.php/cyt/article/download/357/407

15. Galindo, J. (2020). *Inteligencia artificial y medios: renovarse o morir.* https://www.cuadernosdeperiodistas.com/inteligencia-artificial-y-medios-renovarse-o-morir/#:~:text=Seg%C3%BAn%20una%20encuesta%20de%20la,2049)%20y%20trabajando%20como%20cirujano

16. González, C. (2023). *La inteligencia artificial como arma de doble filo: ciberataques sofisticados y sistemas de vanguardia.* https://computerhoy.com/tecnologia/impacto-ia-ciberseguridad-ataques-avanzados-defensas-mejoradas-1241912

17. Guzmán, L., Varela, W., y Briceño, M. (2020). *Ciberseguridad 4.0: Factores que propician el delito de robo de identidad digital por medios informáticos.* https://riico.net/index.php/riico/article/view/1818/1577

18. Idrovo, M. (2020). *Análisis de casos.* https://www.usfq.edu.ec/sites/default/files/2021-01/pea-036-009.pdf

19. James, V. (2018). *Mira a Jordan Peele usar IA para hacer que Barack Obama entregue un anuncio de servicio público sobre noticias falsas.* https://www.theverge.com/tldr/2018/4/17/17247334/ai-fake-news-video-barack-obama-jordan-peele-buzzfeed

20. Kaminski, S. (2023). *Cómo ChatGPT cambiará la ciberseguridad.* https://www.kaspersky.com/blog/chatgpt-cybersecurity/46959/

21. kaspersky. (2023). *¿Qué son los bots? – Definición y Explicación.* https://www.kaspersky.com/resource-center/definitions/what-are-bots

22. kaspersky. (2023). *Cómo identificar noticias falsas.* https://www.kaspersky.com/resource-center/preemptive-safety/how-to-identify-fake-news

23. Lee, J. (2020). *Cómo las noticias falsas afectan las elecciones estadounidenses.* https://www.ucf.edu/news/how-fake-news-affects-u-s-elections/

24. Llano, F. (2022). *Inteligencia Artificial y Filosofia de Derecho.* https://idus.us.es/bitstream/handle/11441/137250/Inteligencia%20artificial_Llano%20Alonso.pdf?sequence=1&isAllowed=y

25. López, P., Lagares, N., y Puentes, I. (2022). *La inteligencia artificial contra la desinformación: una visión desde la comunicación política.* https://www.revistarazonypalabra.org/index.php/ryp/article/view/1891/1620

26. Niss, O. (2023). *La Ciberdefensa ofensiva y la Inteligencia Artificial.* http://edu.ptn.gov.ar/index.php/revistaecae/article/view/235

27. Noreña, D. (2020). *Revisión de literatura.* https://gestion.pe/blog/el-arte-de-emprender-y-fallar/2020/03/revision-de-literatura.html/?ref=gesr

28. Pastor, J. (2023). *Ni Instagram, ni TikTok: ChatGPT ya es la plataforma que más rápido ha crecido en toda la historia de internet.* http://xataka.com/empresas-y-economia/instagram-tiktok-chatgpt-plataforma-que-rapido-ha-crecido-toda-historia-internet

29. Portela, S. (2023). *Panorama de la inteligencia artificial en el dominio de la ciberseguridad.* https://ruidera.uclm.es/xmlui/handle/10578/30853

30. Puig, A. (2023). *Los estafadores usan inteligencia artificial para perfeccionar sus esquemas de emergencia familiar.* https://consumidor.ftc.gov/alertas-para-consumidores/2023/03/los-

estafadores-usan-inteligencia-artificial-para-perfeccionar-sus-esquemas-de-emergencia-familiar

31. Rodrigo , M. D. (2022). *Ciberseguridad en cadenas de suministros inteligentes en América Latina y el Caribe.* https://repositorio.cepal.org/bitstream/handle/11362/48065/1/S2200203_es.pdf

32. Rodríguez, I. (2023). *La inteligencia artificial y el cibercrimen.* https://www.auditool.org/blog/auditoria-de-ti/la-inteligencia-artificial-y-el-cibercrimen

33. Telefonicatech. (2019). *Cómo detectar Fake News con Machine Learning.* https://telefonicatech.com/blog/como-detectar-fake-news-con-machine-learning

34. Visus, A. (2021). *Que es un Deep fakes, cómo se crean, cuáles fueron los primeros y su futuro.* https://www.esic.edu/rethink/tecnologia/deep-fakes-que-es-como-se-crean-primeros-y-futuros#:~:text=Utiliza%20las%20denominadas%20redes%20neuronales,ese%20objeto%2C%20rostro%20o%20imagen.