LA INTELIGENCIA ARTIFICIAL EN LA CIBERSEGURIDAD: PREVENCIÓN Y DETECCIÓN DE AMENAZAS EN LA ERA DIGITAL

Luis Alberto Flores Montaño, Dr.[1], Jacobo Sandoval Gutiérrez, Dr.[1]
[1]Universidad Autónoma Metropolitana, Unidad Lerma
lfloresm1703o@alumno.ipn.mx, j.sandoval@correo.ler.uam.mx

Boletín No. 107, 1o. de marzo de 2025

Resumen

Hoy en día la ciberseguridad es una preocupación creciente en un mundo altamente conectado, donde las ciberamenazas evolucionan constantemente en su complejidad y sofisticación, afectando a diversos sistemas tanto complejos o de bajo procesamiento, afectando a la población en general y a diversas organizaciones. En este artículo se explora como la inteligencia artificial (IA) puede ser funcional en la rama de la ciberseguridad, destacando el potencial y la capacidad para detectar, prevenir y recuperarse de los ataques cibernéticos en tiempo real. Además, se analizan casos prácticos de herramientas que actualmente son utilizadas para la protección de datos utilizando la inteligencias artificial; adicionalmente, se detallan ventajas y desafíos de integrar la IA en sistemas que protejan la seguridad de los dispositivos, adicionalmente, se explora el potencial que se tiene para transformar el panorama de la ciberseguridad si bien la IA ofrece soluciones que pueden ser prometedoras, también plantea algunos retos tanto éticos como técnicos, los cuales deben abordarse para garantizar su eficacia y eficiencia.

Palabras Clave: inteligencia artificial, ciberseguridad, prevención de amenazas, detección de malware, seguridad digital.

1. Introducción

En la época digital actual, los ciberataques se han vuelto más frecuentes y sofisticados, afectando a diversas organizaciones. Desde ataques sencillos como de ataques de fuerza bruta o de ransomware hasta violaciones de datos a gran escala como lo son ataques DDoS o ataques de día cero (Anderson, 2023); todas y cada una de estas ciberamenazas ponen en peligro la integridad, confidencialidad y disponibilidad, así como la privacidad de cada individuo u organizaciones, riesgos en la economía y en la seguridad global. Teniendo en cuenta esta realidad, la IA ha emergido como una herramienta clave para combatir estos riesgos ante ciberataques (Wirkuttis, 2017). Con capacidades como el aprendizaje automático y la detección de patrones anómalos, la IA está revolucionando la forma en que se identifica y previene ciberataques. En este artículo se analizan las aplicaciones que tiene la IA en la ciberseguridad, resaltando la importancia de la detección en tiempo real y la respuesta proactiva a amenazas.

En la figura 1, se muestra algunos de los beneficios y aportes que tiene la IA dentro de la ciberseguridad, este tecnología puede llevar diversos avances como lo es la detección y prevención de fugas de correos como lo es el phishing y el spamming, se pueden prevenir ataques sofisticados como los de día cero, se puede realizar una autenticación más precisa sobre la autorización de los usuarios a ciertas aplicaciones o programas específicos, así mismo se puede hacer la detección de diversos malwares y de vulnerabilidades dentro de un sistema, esto con el fin de tener una seguridad preventiva para las amenazas existentes; adicionalmente, se puede tener un análisis de eventos que se registra en la actividad de los usuarios, y alertar al sistema de monitoreo.

Figura 1 Funciones de la IA dentro de la ciberseguridad (Elaboración propia).

2. Contenido del artículo

A continuación, se describe de manera precisa las funciones de la IA dentro de la ciberseguridad, mostradas en la figura 1.

1. **El papel de la IA en la detección de amenazas:** La IA permite analizar grandes volúmenes de datos en tiempo real, identificando patrones anómalos que podrían indicar actividades maliciosas. Herramientas basadas en IA, como son los sistemas de detección y prevención de intrusiones (por sus siglas en inglés IDS/IPS), emplean algoritmos de aprendizaje automático para diferenciar el tráfico normal del malicioso (Zhang, 2023). Un ejemplo de esto es el uso de redes neuronales profundas para detectar intentos de envío de correos con contenido phishing, esto es prevenido antes de que lleguen al usuario final.
2. **Detección de Email Phishing:** Se puede identificar los correos fraudulentos diseñados para engañar a los usuarios y robar las credenciales o información sensible; con técnicas de inteligencia artificial se puede realizar diversas detecciones para este tipo de correos como lo es el procesamiento de lenguaje natural (NLP), para el análisis de contenido detectando patrones maliciosos, esto puede involucrar errores gramaticales, solicitudes de información o enlaces maliciosos (Cylance, 2023).
3. **Prevención proactiva mediante IA:** Además de detectar amenazas y correos maliciosos, la IA también puede anticiparse a ataques futuros al analizar tendencias históricas y comportamientos sospechosos. Por ejemplo, sistemas de IA pueden identificar vulnerabilidades potenciales en la infraestructura digital y sugerir soluciones antes de que sean explotadas (Darktrace, 2023).
4. **Detección y aplicación de parches automáticos de vulnerabilidades IA:** También la IA puede identificar vulnerabilidades en sistemas y automatiza las descargas e instalación de parches sin intervención humana. Un ejemplo de esto es emplear algoritmos para predecir las vulnerabilidades futuras o priorizar parches acordes con el impacto que se pudiera generar ante ciertas amenazas, con esto reduce la exposición ante ataques del día cero (Darktrace, 2023).
5. **Autenticación de usuario y control de acceso IA:** Además de diversas detecciones y prevenciones que se pueden realizar con ayuda de esta tecnología, también es posible la autenticación para garantizar que

solo los usuarios autorizados accedan a recursos específicos según los permisos. Por ejemplo, la IA puede realizar autenticaciones basadas en el comportamiento, donde algoritmos de redes neuronales pueden detectar patrones como la velocidad de tecleo, movimientos del ratón o interacción con dispositivos para detectar ciertas anomalías (Cylance, 2023).

6. **Análisis de comportamiento de usuario y detección de amenazas IA:** Teniendo en cuenta la detección de patrones, también es posible hacer un análisis y monitoreo completo del comportamiento del usuario para poder detectar actividades anómalas que podrían indicar amenazas internas o externas, un ejemplo de esto es el uso de algoritmos de aprendizaje no supervisado para detectar patrones fuera de lo común (Cylance, 2023).

3. Casos prácticos de IA en ciberseguridad

De lo mencionado anteriormente, se retoman casos prácticos de herramientas que utilizan este tipo de tecnologías dentro de la industria. A continuación, se mencionan algunas de estas.

- **Detección de malware:** Herramientas dentro de la industria como es Cylance utiliza la IA para analizar código de archivos y predecir si son maliciosos, incluso sin firmas conocidas (Cylance, 2023).
- **Respuestas automatizadas:** Sistemas como lo es Darktrace también utilizan la IA para contener ataques en tiempo real, con el propósito de aislar dispositivos comprometidos sin intervención humana (Darktrace, 2023).
- **Protección contra ataques de fuerza bruta:** Como se mencionó anteriormente, la IA puede identificar patrones de intentos de acceso repetidos y bloquearlos automáticamente, dentro de la industria existen diversas herramientas encargadas de detectar ataques de este tipo encargados de la detección temprana, adaptabilidad, reducción de falsos positivos y acciones tempranas, algunas de estas son Fail2Ban con IA, Elastic Security o Azure Sentinel (Pooyandeh, 2022).

Un ejemplo de esto son los algoritmos de aprendizaje automático (machine learning), este tipo de algoritmos son capaces de analizar una gran cantidad de datos en tiempo real, todo esto para identificar ataques que no han sido identificados en una base de datos, esto también es conocido en la industria como ataques de día cero (Zero-Day Attacks). Adicionalmente, las redes neuronales profundas (Deep Learning), en el caso de estos pueden encargarse de clasificar archivos maliciosos y distinguirlos de los benignos, mejorando con más precisión los IPS (Patil, 2016).

4. Respuesta automatizada y mitigación en tiempo real

Por otro lado, la IA ha revolucionado la capacidad de respuesta ante diversos incidentes. Los sistemas basados en IA son capaces de reaccionar automáticamente ante amenazas que estén en curso, con el propósito de bloquear conexiones sospechosas, aislando dispositivos que se encuentran comprometidos (infectados) o en su defecto se pueden añadir parches de seguridad de manera proactiva (Mohammed, 2020).

Los sistemas relacionados con este tipo de tecnología son conocidos como Sistemas de Respuesta Automática a Incidentes (SOAR), estos últimos han ganado popularidad al integrar la IA para coordinar y ejecutar acciones de mitigación en tiempo real, ayudando a reducir significativamente el tiempo de respuesta ante este tipo de incidentes (Das et al., 2021).

5. Predicción de vulnerabilidades y análisis proactivo

Otra de las grandes contribuciones que se tienen de la IA es la capacidad predictiva ante vulnerabilidades; lo anterior es posible debido al análisis de patrones históricos y la correlación entre eventos, así los sistemas de IA pueden predecir posibles ataques antes de que ocurran (Puthal et al., 2021).

Un ejemplo de esto es el caso notable del uso de modelos predictivos en la protección de infraestructuras críticas, como lo son las redes eléctricas y los sistemas de transporte, en las cuales las consecuencias de un ciberataque pueden ser devastador. Estas herramientas también pueden ser empleadas para priorizar las vulnerabilidades en sistemas, debido a lo anterior es posible guiar a los equipos de ciberseguridad hacia áreas de mayor riesgo (Sindiramutty et al., 2024).

6. Herramientas y aplicaciones basadas en IA

Adicionalmente, se tienen diversas herramientas de ciberseguridad integradas con IA. Algunos ejemplos de dichas herramientas son los siguientes:

- **Sistemas de detección de intrusos (IDPS):** Tecnologías como lo es Darktrace utilizan IA para analizar tráfico de las redes y detectar amenazas en tiempo real (Calderon, 2019).
- **Análisis de malware:** Plataformas gratuitas como lo es VirusTotal emplean modelos de IA para identificar y clasificar malware, esta herramienta lo permite hacer a través de diversos archivos, así como el análisis de diversos enlaces (Wirkuttis, 2017).
- **Autenticación biométrica:** Estos pueden ser sistemas basados en reconocimiento facial y de huellas digitales respaldos por los algoritmos basados en IA (Das et al., 2021).

7. Retos y limitaciones de la IA en ciberseguridad

No obstante, a pesar de los distintos avances de esta tecnología, no está exenta a diversos desafíos. Algunos de estos desafíos incluyen los siguientes:

- **Falsos positivos y negativos:** Aunque la IA ha mejorado bastante en la detección de malwares o ataques, los errores en los análisis pueden causar alertas erróneas o en su defecto innecesarias, y por otro lado, pueden pasar por alto amenazas reales (Ansari et al., 2022).
- **Dependencia de datos:** Los modelos de IA requieren de un vasto volumen de datos para que puedan entrenarse, lo que significa un gran obstáculo en los sectores, en donde los datos pueden ser limitados o sensibles (Familoni, 2024).
- **Uso malicioso de la IA:** La otra cara de la moneda en el uso de la IA es en el uso indebido donde emplearla es posible desarrollar ciberataques y malware más sofisticado, y así evadir los sistemas de ciberseguridad (Yampolskiy, 2016).

8. Conclusiones

La inteligencia artificial ha transformado en gran parte la ciberseguridad, lo cual ha permitido a las organizaciones en responder a las amenazas de forma más precisa, rápida y proactiva. Ya que al aprovechar la capacidad de la IA es capaz de analizar datos de grandes volúmenes, predecir ataques y automatizar respuestas, lo que ha alcanzado los niveles de protección sin precedentes.

No obstante, esta herramienta tecnológica también puede plantear nuevos desafíos como lo es la necesidad de manejar datos de forma responsable, y prevenir su mal uso. Por lo que a medida que la IA va evolucionando, la integración con la ciberseguridad es cada vez más esencial para mantenerse al margen de las ciberamenazas en un ciberespacio que cada vez crece más día a día.

Referencias

[1] Ansari, M. F., Dash, B., Sharma, P., & Yathiraju, N. (**2022**). *The impact and limitations of artificial intelligence in cybersecurity: a literature review.* International Journal of Advanced Research in Computer and Communication Engineering.

[2] Calderon, R. (**2019**). *The benefits of artificial intelligence in cybersecurity.*

[3] Das, R., & Sandhane, R. (**2021, July**). *Artificial intelligence in cyber security.* In Journal of Physics: Conference Series (Vol. 1964, No. 4, p. 042072). IOP Publishing.

[4] Familoni, B. T. (**2024**). *Cybersecurity challenges in the age of AI: theoretical approaches and practical solutions.* Computer Science & IT Research Journal 5(3), 703-724.

[5] Patil, P. (**2016**). *Artificial intelligence in cybersecurity.* International journal of research in computer applications and robotics, 4(5), 1-5.

[6] Puthal, D., & Mohanty, S. P. (**2021**). *Cybersecurity issues in AI.* IEEE Consumer Electronics Magazine, 10(4), 33-35.

[7] Mohammed, I. A. (**2020**). *Artificial intelligence for cybersecurity: A systematic mapping of literature.* Artif. Intell 7(9), 1-5.

[8] Sindiramutty, S. R., Tan, C. E., Lau, S. P., Thangaveloo, R., Gharib, A. H., Manchuri, A. R., ... & Muniandy, L. (**2024**). *Explainable AI for Cybersecurity.* In Advances in Explainable AI Applications for Smart Cities (pp. 31-97). IGI Global.

[9] Yampolskiy, R. V., & Spellchecker, M. S. (**2016**). *Artificial intelligence safety and cybersecurity: A timeline of AI failures.* arXiv preprint arXiv:1610.07997.

[10] Wirkuttis, N., & Klein, H. (**2017**). *Artificial intelligence in cybersecurity.* Cyber, Intelligence, and Security, 1(1), 103-119.

Flores Montaño, L. A., Sandoval Gutiérrez, J. *(2026). LA INTELIGENCIA ARTIFICIAL EN LA CIBERSEGURIDAD: PREVENCIÓN Y DETECCIÓN DE AMENAZAS EN LA ERA DIGITAL. Boletín UPIITA. año XX, (NÚM) 2026.*