

Una nueva ola de estafas impulsadas por la IA

El avance de la Inteligencia Artificial (IA) ha transformado para siempre nuestra vida digital. Aunque los chatbots y algoritmos hacen más fácil y eficiente nuestra experiencia en línea, también han creado nuevas amenazas de ingeniería social, que al igual que las estafas tradicionales, buscan robar datos personales, información bancaria y detalles sensibles, tanto de personas como de empresas.

Con la IA, los ciberdelincuentes han perfeccionado tácticas como las estafas de phishing, cuyos ataques en Chile aumentaron 125% en 2024, en comparación con el año anterior, según el Panorama de Amenazas de Kaspersky. A través de mensajes falsos, los criminales engañan a sus víctimas para que revelen inadvertidamente sus correos, contraseñas o datos de sus tarjetas bancarias. Se dirigen tanto a usuarios como a empresas, de forma masiva o personalizada, disfrazados de notificaciones de bancos, proveedores de servicios, sistemas de pago electrónico u otras organizaciones, incluso puede parecer que provienen de alguien que conocemos.

Otro ejemplo son las ultrafalsificaciones, también conocidas como deepfakes. Con apenas unos segundos de una grabación de



Si bien los deepfakes y las estafas impulsadas por la IA presentan retos cada vez mayores, comprender estas amenazas es el primer paso para enfrentarlas. No hay por qué tener miedo, si combinamos soluciones y mejores prácticas de seguridad con una adecuada alfabetización cibernética, tanto usuarios como organizaciones podemos reducir los riesgos y contribuir a la construcción de un entorno digital más seguro y resiliente.

voz, la IA puede generar clips de audio; además, también facilita la alteración de imágenes y videos modificando rostros y sus expresiones. Se estima que para 2026, hasta el 90% del contenido en línea podría generarse de esta forma. Esto es preocupante porque el riesgo de suplantación de identidad aumenta.

Los cibercriminales han adoptado estas herramientas avanzadas para emplear tácticas más complejas. Por ejemplo, roban cuentas de aplicaciones de mensajería, como Telegram o WhatsApp; con los mensajes de voz en los chats crean grabaciones que imitan a los dueños de esas cuentas y las envían a contactos de confianza de las víctimas, como amigos, familiares o clientes, para estafarlos.

En resumen, con la Inteligencia Artificial los criminales pueden automatizar la producción masiva de contenido fraudulento, haciendo sus ataques más sofisticados y difíciles de detectar. Por eso, a medida que esta tecnología avanza, nuestra defensa debe enfocarse en dos frentes: técnico y educativo.

A nivel técnico, existen soluciones prometedoras que se pueden adoptar, como las marcas de agua, para etiquetar contenido generado por IA; detectores de deepfakes,

para identificar características específicas de contenido manipulado; y firmas digitales, las cuales ya se utilizan en transacciones bancarias y comunicaciones importantes, para verificar la autenticidad de imágenes, videos y audios alterados. El principal desafío de estas medidas es evolucionar a la misma velocidad que los grandes modelos de lenguaje y la IA generativa, exigiendo una constante actualización para que sean efectivas.

A nivel educativo, hay una brecha crítica: un desconocimiento de lo fácil que es explotar la Inteligencia Artificial. La ciberdelincuencia aprovecha esta falta de conocimiento, lo que resalta la necesidad de un diálogo abierto y de campañas educativas sobre los riesgos.

Si bien los deepfakes y las estafas impulsadas por la IA presentan retos cada vez mayores, comprender estas amenazas es el primer paso para enfrentarlas. No hay por qué tener miedo, si combinamos soluciones y mejores prácticas de seguridad con una adecuada alfabetización cibernética, tanto usuarios como organizaciones podemos reducir los riesgos y contribuir a la construcción de un entorno digital más seguro y resiliente.



ISABEL MANJARREZ
 INVESTIGADORA DE SEGURIDAD
 DEL EQUIPO GLOBAL DE ANÁLISIS
 DE KASPERSKY