Durante la charla,el Dr. van Ede explicó cómo la inteligencia artificial puede contribuir a automatizar los procesos de defensa digital ante los crecientes ataques informáticos.Foto cortesía de Herson Esquivel /TEC.

La inteligencia artificial entra en el campo de batalla digital

Inteligencia artificial y ciberseguridad: investigador de Twente presenta “Terminator”, un modelo para mitigar ciberataques

21 de Octubre 2025 Por: Irina Grajales Navarrete [1]

Durante su visita al TEC, el Dr. Thijs van Ede presentó “Terminator”, un proyecto que busca entender cómo la inteligencia artificial puede automatizar la defensa cibernética sin caer en los peligros de una inteligencia fuera de control

La Unidad de Posgrado en Computación [2] del Tecnológico de Costa Rica (TEC) [3] organizó la charla **“Terminator: usar la IA para mitigar ciberataques”**, impartida por el Dr. Thijs van Ede,

profesor asistente de la Universidad de Twente, Países Bajos, especialista en inteligencia artificial y ciberseguridad. La actividad se realizó en modalidad híbrida desde el recinto Zapote del Campus Tecnológico Local San José [4].

Durante su intervención, **el Dr. van Ede explicó cómo la inteligencia artificial puede contribuir a automatizar los procesos de defensa digital ante los crecientes ataques informáticos**, un campo que hoy en día se enfrenta a desafíos éticos y técnicos de gran complejidad.

> “El objetivo es investigar cómo la IA puede ayudar a responder de forma más rápida y eficaz ante amenazas, sin poner en riesgo la seguridad del sistema mismo”, señaló el especialista.

El investigador inició su charla con una breve introducción sobre la Universidad de Twente, ubicada en el este de los Países Bajos, cerca de la frontera con Alemania, y presentó al grupo de investigación en el que participa, conformado por especialistas en ciencia de datos, inteligencia artificial y ciberseguridad.

Posteriormente, explicó **cómo funcionan las operaciones de seguridad dentro de una red de computadoras**. Mediante el uso de agentes de monitoreo y sistemas de detección de intrusos, las instituciones recopilan información sobre actividades sospechosas. Estos sistemas, dijo, utilizan reglas y modelos de aprendizaje automático para detectar comportamientos anómalos, como accesos indebidos a archivos sensibles o intentos de ingreso fallidos.

“En la actualidad, la mayoría de las tareas en los centros de operaciones de

seguridad se realizan de manera manual. Nuestro propósito es explorar cómo la IA puede automatizar estos procesos sin comprometer la seguridad", añadió.

El proyecto "Terminator" —nombre que, según el ponente, **busca generar conciencia sobre los posibles riesgos de una inteligencia artificial mal utilizada**— plantea límites éticos y técnicos en el uso de la automatización para la defensa cibernética.

"En la película Terminator, la IA se vuelve contra la humanidad. Nosotros queremos exactamente lo contrario: investigar cómo evitar ese escenario y entender las

plicó el Dr.

Durante la sesión, varios asistentes participaron con preguntas sobre los posibles riesgos y reacciones negativas ante este tipo de desarrollos. El ponente reconoció que los peligros son significativos si un modelo de IA llegara a ejecutar acciones como eliminar archivos del sistema o alterar redes críticas.

"El potencial de consecuencias negativas es enorme, y precisamente por eso debemos investigarlo. Si no lo hacemos, podríamos enfrentarnos a escenarios indeseados", advirtió.

El profesor destacó también la asimetría entre atacantes y defensores en el ciberespacio: mientras los ataques ya se automatizan, los defensores deben ser más cautelosos, pues un error en la defensa puede resultar desastroso.

Finalmente, respondió consultas sobre los métodos de validación de modelos de IA, señalando que su equipo basa los casos de prueba en el Migrate Defense Framework, un marco que clasifica las técnicas de defensa en categorías como aislar (isolate), expulsar (evict) y restaurar (restore).

“No existe aún un estándar, pero estamos trabajando en ello. Este es solo un primer paso hacia un modelo más seguro y transparente para el uso de IA en ciberseguridad”, concluyó.

Enlaces
[1] https://www.tec.ac.cr/hoyeneltec/users/irina-grajales-navarrete
[2] https://www.tec.ac.cr/unidad-posgrados-computacion
[3] https://www.tec.ac.cr/
[4] https://www.tec.ac.cr/campus-tecnologico-local-san-jose