La inteligencia artificial plantea peligros por uso indebido, dicen investigadores

24/02/2018

Un nuevo estudio, publicado esta semana por 25 investigadores técnicos y de políticas públicas de las universidades de Cambridge, Oxford y Yale junto con expertos en privacidad y militares, hizo sonar las alarmas por el posible uso indebido de la IA por parte de estados enemigos, delincuentes y lobos solitarios.

Los investigadores dijeron que el uso malicioso de la IA representa una amenaza inminente para la seguridad digital, física y política al posibilitar ataques a gran escala y altamente eficaces. El estudio se centra en desarrollos plausibles dentro de cinco años.

“Todos estamos de acuerdo en que hay muchas aplicaciones positivas para la IA”, dijo Miles Brundage, investigador del Instituto Future of Humanity de Oxford. “Hubo un blanco en la literatura sobre el tema del uso malicioso”.

La inteligencia artificial, o IA, implica el uso de ordenadores para realizar tareas que normalmente requieren inteligencia humana, como tomar decisiones o reconocer texto, voz o imágenes visuales.

Se considera una poderosa fuerza para desbloquear todo tipo de posibilidades técnicas, pero se ha convertido en el centro de acalorados debates sobre si la automatización masiva que permite podría resultar en un desempleo generalizado y otras perturbaciones sociales.

El documento de 98 páginas advierte que el coste de los ataques puede reducirse mediante el uso de la IA para completar tareas que de otra manera requerirían trabajo humano y experiencia. Pueden producirse nuevas formas de ataques que no podrían lanzar los humanos sin ayuda de la IA o que explotan las vulnerabilidades de los propios sistemas de inteligencia artificial.

El informe revisa un creciente material de investigación académica sobre los riesgos de seguridad que plantea la IA y hace un llamamiento a los gobiernos y expertos en políticas y técnicos para que colaboren y atenúen estos peligros.

Los investigadores detallan el poder de la IA para generar imágenes sintéticas, texto y audio para suplantar a otros en la red, con el fin de influir en la opinión pública, señalando la amenaza de que gobiernos autoritarios usen dicha tecnología.

El informe hace una serie de recomendaciones que incluyen la regulación de la inteligencia artificial como una tecnología militar/comercial con un doble uso.

También cuestiona sobre si los académicos y otros deben controlar lo que publican o divulgan sobre los nuevos progresos en IA hasta que otros expertos en el campo tengan la oportunidad de estudiar y reaccionar ante los peligros potenciales que puedan plantear.

“Al final, terminamos con muchas más preguntas que respuestas”, dijo Brundage.

El documento nació de un seminario a principios de 2017, y algunas de sus predicciones se cumplieron mientras se escribía. Por ejemplo, los autores especularon que la IA se podría usar para crear audio y video falsos sumamente realistas de autoridades públicas con fines de propaganda.