ACTUALIDAD / CIENCIA Y TECNOLOGÍA

Inteligencia artificial podría ser capaz de identificar cuando un resultado no es confiable

El Ciudadano · 26 de noviembre de 2020

Los investigadores estiman que conocer el nivel de incertidumbre en las decisiones tomadas por esos sistemas podría brindar resultados más seguros

Un grupo de investigadores del Instituto de Tecnología de Massachusetts (MIT, por sus siglas en inglés) y de la Universidad de Harvard, ambas de EE.UU., desarrollaron un sistema para saber si las conclusiones de los sistemas de inteligencia artificial conocidos como redes neuronales de aprendizaje profundo son correctas o no. La utilidad del proceso radica en que este tipo de redes son cada vez más utilizadas en la vida cotidiana, como en la conducción de un vehículo autónomo o en la medicina.

Para ello, elaboraron una forma rápida para que la red no solamente genere una predicción, sino que también determine el nivel de confianza en los resultados, en base a la calidad de los datos con los que se cuenta. Para ello, Alexander Amini, estudiante de doctorado en el Laboratorio de Ciencias de la Computación e Inteligencia Artificial del MIT (CSAIL, por sus siglas en inglés), recurrió a un enfoque llamado 'regresión probatoria profunda', con el que se podría llegar a resultados más seguros.

«Necesitamos la capacidad no solo de tener modelos de alto rendimiento, sino también de comprender cuándo no podemos confiar en esos modelos», afirmó Amini. Por su parte, Daniela Rus, profesora del CSAIL, agregó: «Esta idea es importante y aplicable ampliamente. Se puede utilizar para evaluar productos que se basan en modelos aprendidos. Al estimar su incertidumbre, también aprendemos cuánto error esperar del modelo y qué datos faltantes podrían mejorarlo».

Efectividad del 99 %

En cuanto a la efectividad que tienen los procesos de las redes neuronales, Amini aseguró que «son realmente buenas para saber la respuesta correcta el 99 % de las veces». Sin embargo, manifestó su preocupación «por ese 1 %», por lo que es necesario «detectar esas situaciones de manera confiable y eficiente».

Para ello, desarrollaron una manera que permite a la red tomar una decisión y reflejar las evidencias que la apoyan. De esta manera, se puede determinar si esa incertidumbre puede ser reducida o si la información con la que trabajó la red no es precisa.

Este procedimiento fue probado a través del análisis de una imagen en color monocular y estimaron un valor de profundidad (distancia desde la lente de la cámara) para cada píxel. Este tipo de cálculos son los que podría utilizar un vehículo de conducción autónoma para determinar su cercanía con otro vehículo o con un peatón.

La red proyectó un alto nivel de incertidumbre referida a los píxeles en los que la profundidad que predijo fue incorrecta. «Estaba muy calibrado para los errores que comete la red, que creemos que fue una de las cosas más importantes para juzgar la calidad de un nuevo estimador de incertidumbre», explicó Amini.

«Cualquier usuario del método, ya sea un médico o una persona en el asiento del pasajero de un vehículo, debe ser consciente de cualquier riesgo o incertidumbre» asociados con las decisiones de la red señaló Amini, quien estima que el sistema también utilizará el resultado para tomar decisiones que eviten el peligro. «Cualquier campo que vaya a tener aprendizaje automático, en última instancia, debe tener un conocimiento confiable de la incertidumbre», concluyó.

Cortesía de RT

Te podría interesar

Inteligencia artificial revela emociones en aula virtual

Fuente: El Ciudadano