

## Padrino de la inteligencia artificial que renunció a Google gana Nobel de Física

► Geoffrey Hinton, académico de la Universidad de Toronto, ha manifestado sus miedos en torno al desarrollo de la inteligencia artificial.

**Geoffrey Hinton** se adjudicó el premio junto a John Hopfield, quienes utilizaron herramientas que ayudaron a sentar las bases del poderoso aprendizaje automático actual, y que se ha vuelto importante para las grandes redes neuronales artificiales que se utilizan actualmente.

**Francisco Corvalán**

El Premio Nobel de Física de este año fue otorgado a los académicos de las universidades de Princeton y Toronto, John J. Hopfield y Geoffrey E. Hinton, respectivamente. "El trabajo de los galardonados ya ha sido de gran beneficio. En física utilizamos redes neuronales artificiales en una amplia gama de áreas, como el desarrollo de nuevos materiales con propiedades específicas", dijo Ellen Moons, presidenta del Comité Nobel de Física.

"Ambos han sido verdaderos pioneros: encontraron nuevas formas de abordar los problemas", declaró también el profesor Anders Irbäck, miembro del Comité Nobel. Por su parte, la noticia encontró a Hinton totalmente desprevenido. "Estoy en un hotel barato en California que no tiene buena conexión a Internet ni a teléfono. Iba a hacerme una resonancia magnética hoy, pero tendré que cancelarla", manifestó el llama-

do "padrino de la IA" luego de enterarse del anuncio.

### Las máquinas "aprenden"

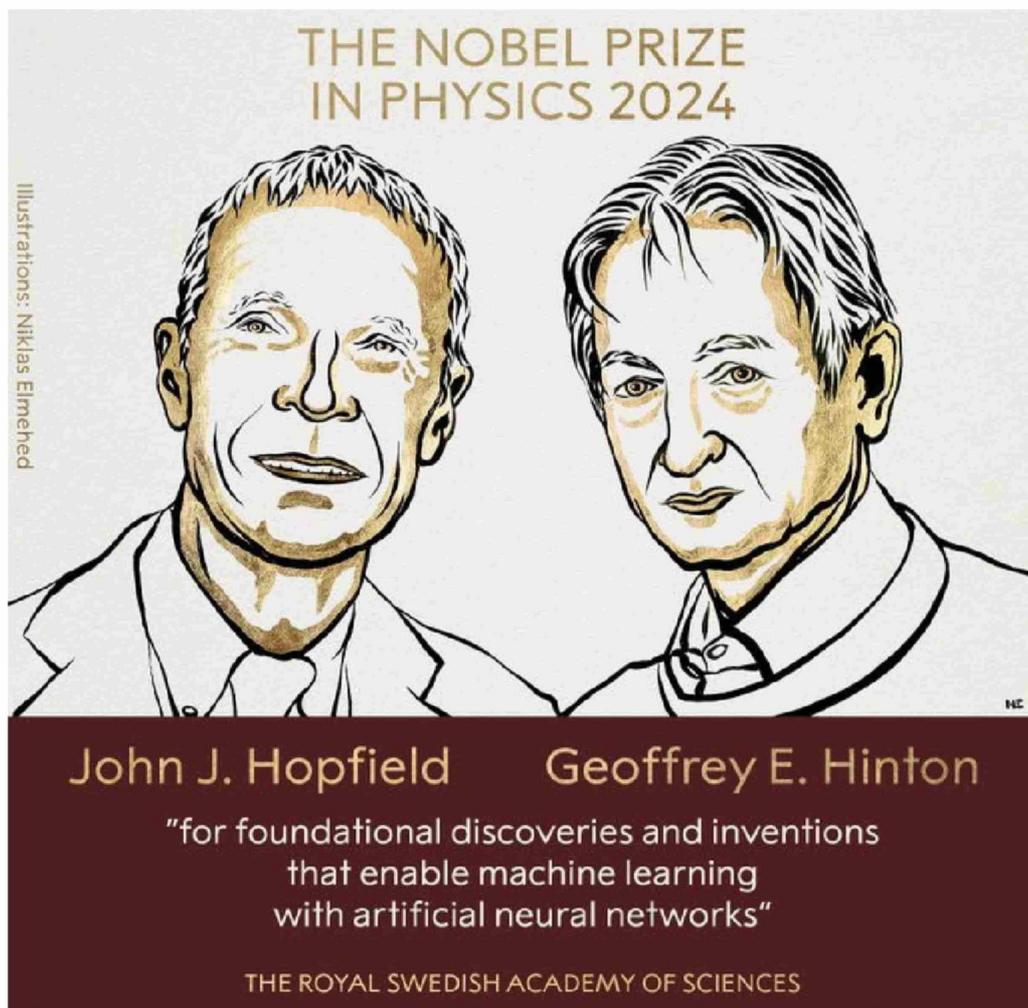
Además, confesó estar atónito con la noticia, "no tenía idea de que esto sucedería. Estoy muy sorprendido", dijo. Hinton, que ha hablado sobre sus miedos en torno al desarrollo de la inteligencia artificial, expresó que "en las mismas circunstancias, volvería a hacer lo mismo, pero me preocupa que las consecuencias generales de esto puedan ser sistemas más inteligentes que nosotros que eventualmente tomen el control".

De hecho, en mayo de 2023 Hinton renunció a su cargo en Google para advertir de la "amenaza existencial" que representa la inteligencia artificial. En ese momento, el Centro para la Seguridad de la IA publicó una carta abierta, firmada por Hinton y cientos de otras personas, en la que advertía de que la Inteligencia Artificial avanzada podría destruir a la humanidad.

Pero, ¿por qué esta dupla de padrinos de la IA obtuvieron el Nobel? Cuando hablamos de inteligencia artificial, a menudo nos referimos al aprendizaje automático mediante redes neuronales artificiales. Esta tecnología se inspiró originalmente en la estructura del cerebro. En una red neuronal artificial, las neuronas del cerebro están representadas por nodos que tienen diferentes valores.

Estos nodos se influyen mutuamente a través de conexiones que pueden compararse con las sinapsis y que pueden fortalecerse o debilitarse. La red se entrena, por ejemplo, mediante el desarrollo de conexiones más fuertes entre nodos con valores simultáneamente altos. Los galardonados de este año han llevado a cabo un importante trabajo con redes neuronales artificiales desde la década de 1980 en adelante.

SIGUE ►►



► Los ganadores del Premio Nobel de Física 2024.

**SIGUE ►►**

Muchas personas han experimentado cómo las computadoras pueden traducir entre idiomas, interpretar imágenes e incluso mantener conversaciones razonables. Lo que quizás sea menos conocido es que este tipo de tecnología ha sido importante durante mucho tiempo para la investigación, incluida la clasificación y el análisis de grandes cantidades de datos. El desarrollo del aprendizaje automático se ha disparado en los últimos quince o veinte años y utiliza una estructura llamada red neuronal artificial. Hoy en día, cuando hablamos de inteligencia artificial, a menudo nos referimos a este tipo de tecnología.

Aunque las computadoras no pueden pensar, ahora pueden imitar funciones como la memoria y el aprendizaje. Los laureados de este año en física ayudaron a hacer esto posible. Utilizando conceptos y métodos fundamentales de la física, han desarrolla-

do tecnologías que utilizan estructuras en redes para procesar información.

El aprendizaje automático difiere del software tradicional, que funciona como un tipo de receta. El software recibe datos, que se procesan de acuerdo con una descripción clara y producen los resultados, al igual que cuando alguien recolecta ingredientes y los procesa siguiendo una receta, produciendo un pastel. En lugar de esto, en el aprendizaje automático la computadora aprende con el ejemplo, lo que le permite abordar problemas que son demasiado vagos y complicados para ser manejados por instrucciones paso a paso. Un ejemplo es la interpretación de una imagen para identificar los objetos que contiene.

Una red neuronal artificial procesa la información utilizando toda la estructura de la red. Inicialmente, la inspiración vino del deseo de entender cómo funciona el cerebro. En la década de 1940, los investigado-

res habían comenzado a razonar en torno a las matemáticas que subyacen a la red de neuronas y sinapsis del cerebro. Otra pieza del rompecabezas vino de la psicología, gracias a la hipótesis del neurocientífico Donald Hebb sobre cómo se produce el aprendizaje porque las conexiones entre las neuronas se refuerzan cuando trabajan juntas.

Más tarde, estas ideas fueron seguidas por intentos de recrear cómo funciona la red del cerebro mediante la construcción de redes neuronales artificiales como simulaciones por computadora. En estos, las neuronas del cerebro son imitadas por nodos a los que se les dan diferentes valores, y las sinapsis están representadas por conexiones entre los nodos que se pueden hacer más fuertes o más débiles. La hipótesis de Donald Hebb todavía se utiliza como una de las reglas básicas para actualizar redes artificiales a través de un proceso llamado entrenamiento.

A finales de la década de 1960, algunos resultados teóricos desalentadores hicieron

que muchos investigadores sospecharan que estas redes neuronales no tendrían ninguna utilidad real. Sin embargo, el interés por las redes neuronales artificiales se reavivó en la década de 1980, cuando varias ideas importantes tuvieron impacto, incluido el trabajo de los galardonados de este año.

John Hopfield inventó una red que utiliza un método para guardar y recrear patrones. Podemos imaginar los nodos como píxeles. La red de Hopfield utiliza la física que describe las características de un material debido a su espín atómico, una propiedad que hace que cada átomo sea un pequeño imán.

La red en su conjunto se describe de una manera equivalente a la energía en el sistema de espín que se encuentra en física, y se entrena encontrando valores para las conexiones entre los nodos de modo que las imágenes guardadas tengan baja energía. Cuando la red de Hopfield recibe una imagen distorsionada o incompleta, trabaja metódicamente a través de los nodos y actualiza sus valores para que la energía de la red disminuya. Por lo tanto, la red trabaja paso a paso para encontrar la imagen guardada que más se parece a la imperfecta con la que se alimentó.

Esta red que construyó Hopfield tiene nodos que están unidos entre sí a través de conexiones de diferentes fuerzas. Cada nodo puede almacenar un valor individual: en el primer trabajo de Hopfield, podría ser 0 o 1, como los píxeles de una imagen en blanco y negro. La red Hopfield se puede utilizar para recrear datos que contienen ruido o que se han borrado parcialmente.

Cuando Hopfield publicó su artículo sobre la memoria asociativa, Geoffrey Hinton trabajaba en la Universidad Carnegie Mellon en Pittsburgh, Estados Unidos. Anteriormente había estudiado psicología experimental e inteligencia artificial en Inglaterra y Escocia y se preguntaba si las máquinas podrían aprender a procesar patrones de manera similar a los humanos, encontrando sus propias categorías para clasificar e interpretar la información. Junto con su colega, Terrence Sejnowski, Hinton partió de la red de Hopfield y la amplió para construir algo nuevo, utilizando ideas de la física estadística.

Fue así como Geoffrey Hinton utilizó la red de Hopfield como base para una nueva red que utiliza un método diferente: la máquina de Boltzmann. Esto puede aprender a reconocer elementos característicos en un tipo determinado de datos. Hinton utilizó herramientas de la física estadística, la ciencia de los sistemas construidos a partir de muchos componentes similares.

La máquina se entrena alimentándola con ejemplos que es muy probable que surjan cuando la máquina esté en funcionamiento. La máquina de Boltzmann se puede utilizar para clasificar imágenes o crear nuevos ejemplos del tipo de patrón en el que se entrenó. Hinton se ha basado en este trabajo, ayudando a iniciar el explosivo desarrollo actual del aprendizaje automático. ●