

Una conexión inesperada entre la física de partículas y la inteligencia artificial

Un nuevo resultado permite usar la teoría cuántica de campos para describir algoritmos de aprendizaje automático. El hallazgo ofrece una nueva vía para entender la inteligencia artificial, así como para investigar los fundamentos matemáticos de la propia teoría de campos

DIMITRIOS BACHTIS



LA TEORÍA DE GRAFOS ha permitido establecer una equivalencia entre ciertos algoritmos de aprendizaje automático basados en redes neuronales y la teoría cuántica de campos, el formalismo usado por los físicos para describir las partículas elementales.

La teoría cuántica de campos es el formalismo empleado por los físicos para describir el comportamiento de las partículas elementales, como electrones, fotones o quarks. Hasta hoy, esta teoría ha demostrado tener un éxito rotundo. Por recordar un ejemplo célebre, el cálculo del momento magnético del electrón ha conducido a la verificación experimental más precisa de una predicción teórica en toda la historia de la ciencia.

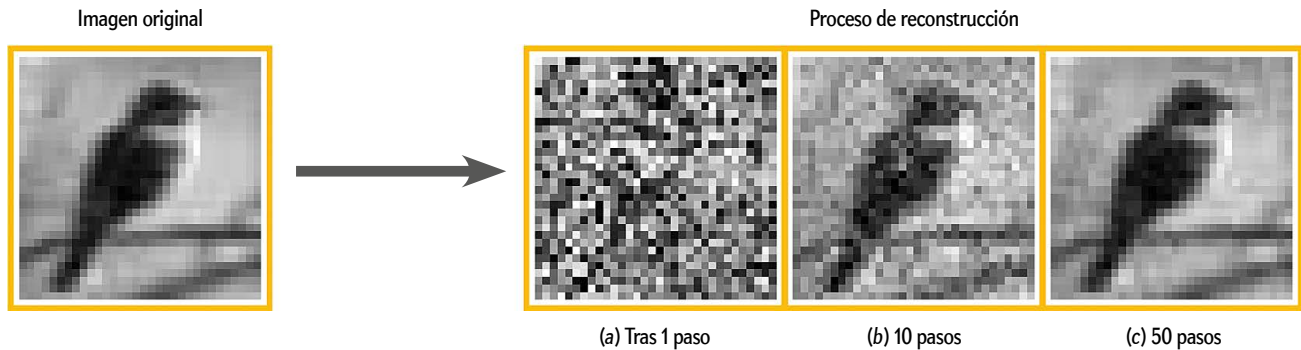
Desde sus inicios hace unos 90 años, la teoría cuántica de campos fue concebida para dar cuenta de la naturaleza a su nivel más fundamental: el de las partículas elementales y sus campos cuánticos correspondientes. Pero ¿podría este formalismo

resultar igualmente exitoso a la hora de describir un comportamiento «emergente»; es decir, el tipo de fenómenos colectivos que surgen cuando muchos constituyentes básicos interactúan entre sí?

Un ejemplo típico de esta clase de comportamiento nos lo proporciona el cerebro humano. Aunque este se compone de elementos simples (neuronas) que interactúan entre sí mediante reglas relativamente sencillas, esas interacciones dan rápidamente lugar a fenómenos colectivos genuinamente nuevos. Estos surgen en última instancia de las interacciones entre neuronas, pero afectan al cerebro como un todo. Por tanto, podemos preguntarnos si es posible concebir una teo-

ría física que describa con precisión esta clase de fenómenos colectivos.

En un trabajo reciente realizado por nuestro grupo de la Universidad de Swansea y publicado el pasado mes de abril en *Physical Review D*, hemos descubierto una sorprendente relación entre ciertas teorías cuánticas de campos y algunos algoritmos de aprendizaje automático. En concreto, hemos demostrado que algunas redes neuronales y otros sistemas de inteligencia artificial pueden describirse directamente mediante modelos computacionales de teorías cuánticas de campos. Este inesperado resultado allana el camino para estudiar el aprendizaje automático mediante el formalismo de la



REINTERPRETAR EL APRENDIZAJE: Este algoritmo es capaz de aprender una imagen dada (izquierda) y, posteriormente, comenzar con un conjunto aleatorio de píxeles (derecha, a), reordenarlos y construir a partir de ellos una versión fiel de la imagen original (derecha, c). Es posible demostrar que este algoritmo es equivalente a una teoría cuántica de campos. En términos de esta última, todo el proceso algorítmico puede reinterpretarse como un sistema físico que alcanza una configuración de equilibrio.

teoría de campos, al tiempo que proporciona una nueva y prometedora vía para investigar los fundamentos matemáticos de la propia teoría de campos.

Espaciotiempo y superordenadores

La conexión entre la teoría de campos y la inteligencia artificial tiene su origen en cierto formalismo usado en física de partículas y conocido como «teoría de campos en el retículo» (*lattice field theory*). Este marco se desvía del uso mayoritariamente en física de partículas, y su característica principal reside en que modeliza el espacio y el tiempo como si estuvieran formados por «píxeles» de tamaño finito. Esta forma de tratar el espaciotiempo tiene la ventaja de que logra mantener bajo control las ecuaciones que deben resolverse. Sin embargo, exige resolver un número gigantesco de ellas, lo que implica que solo puede llevarse a cabo con la ayuda de potentes superordenadores.

A pesar de que esta técnica computacional en física de partículas se remonta a la década de 1970, ha sido solo en los últimos años, gracias a la llegada de los superordenadores modernos, cuando ha conseguido augmentar su precisión. En la búsqueda de métodos para acelerar aún más estos exigentes cálculos, el aprendizaje automático ha sido propuesto como una de las posibles soluciones.

Uno de los métodos de aprendizaje automático más usados en la teoría de campos en el retículo es el basado en redes neuronales. Una red neuronal no es más que un conjunto de nodos («neuronas») interconectados por enlaces. Desde un punto de vista matemático, tales redes pueden describirse mediante un grafo: la representación abstracta de un conjunto de vértices conectados entre sí por aris-

tas. Si identificamos cada vértice con una variable de cierto tipo, las aristas que conectan distintos vértices introducen una forma de dependencia entre las variables correspondientes.

Un aspecto clave es que tales estructuras basadas en grafos se usan también en la teoría de campos en el retículo; en concreto, para representar un campo cuántico en un espaciotiempo discreto, o «pixelado». Gracias a este paralelismo, resulta posible encontrar un lenguaje matemático común para describir tanto el aprendizaje automático como la teoría de campos en el retículo.

Localidad y amnesia

Dicho lenguaje matemático común se conoce en términos técnicos con el nombre de «campos aleatorios de Márkov». De manera simplificada, este formalismo describe un conjunto de variables aleatorias, cada una de las cuales puede identificarse con los nodos de un grafo. Tales variables han de satisfacer la llamada «propiedad de Márkov», así denominada en honor al matemático ruso del siglo XIX Andréi Márkov. Esta propiedad establece que los eventos que ocurren en una determinada región del grafo han de ser independientes de los que suceden en las zonas alejadas.

La propiedad de Márkov está relacionada con una característica conocida como «amnesia» (*memorylessness*): el hecho de que el estado de un sistema solo depende de lo que ha ocurrido en el instante anterior, pero no de aquellos acontecimientos que sucedieron más lejos en el pasado. En algunos algoritmos de aprendizaje automático, como los usados en el tratamiento de imágenes, esta propiedad de Márkov se emplea para descubrir estructuras locales en las imágenes.

En el contexto de los campos cuánticos, la condición de Márkov puede aparecer debido a la discretización del espaciotiempo. En este caso, la propiedad de Márkov equivale a la condición de localidad: el conocido principio físico que impone que los sucesos que ocurren en una determinada región del espaciotiempo solo pueden verse afectados por los acontecimientos cercanos, pero no por aquellos que suceden en puntos lejanos (es decir; el mismo principio que prohíbe la existencia de acciones instantáneas entre puntos distantes del espaciotiempo).

Todos estos paralelismos permiten anticipar una relación entre la teoría cuántica de campos y el aprendizaje automático. En nuestro trabajo, hemos establecido que este es efectivamente el caso al demostrar que ciertas teorías de campos satisfacen cierto teorema conocido como «teorema de Hammersley-Clifford». Este garantiza rigurosamente que dichas teorías de campos satisfacen la condición de Márkov y que, por tanto, pueden reformularse en términos de algoritmos de aprendizaje automático.

Campos cuánticos que aprenden

Nuestro resultado abre la posibilidad de emplear las teorías de campos para derivar nuevos algoritmos de aprendizaje automático para tareas concretas, como la segmentación de imágenes (el proceso de dividir una imagen en diferentes partes, como las áreas claras y oscuras de una fotografía en blanco y negro, por ejemplo). Esto es posible porque las propiedades físicas de los campos cuánticos, como su tendencia a minimizar la energía y otras cantidades, se corresponden con los procesos de optimización usados en el aprendizaje automático.

Por ejemplo, en nuestro trabajo hemos demostrado la equivalencia entre cierta teoría cuántica de campos y un algoritmo al que es posible entrenar para que aprenda una imagen. Tras el entrenamiento, si al algoritmo se le da un conjunto aleatorio de píxeles, los reordenará hasta llegar a una «configuración de equilibrio» que reproduce fielmente la imagen original.

Esta equivalencia entre algoritmos y teorías de campos podría dar lugar a nuevas ideas en inteligencia artificial. Por ejemplo, hace tiempo que se sabe que algunos algoritmos pueden relacionarse con la descripción matemática de ciertos sistemas físicos, como los llamados «vidrios de espín» (una línea de investigación que inició el físico italiano Giorgio Parisi y que este año se ha visto reconocida con la concesión del premio Nobel de física). Nuestros resultados podrían arrojar luz sobre cómo interpretar esos algoritmos de aprendizaje automático, los cuales pueden ahora investigarse con las herramientas propias de la teoría de campos.

Al respecto, sería interesante explorar las posibles transiciones de fase en aquellos algoritmos que tienen un equivalente en teoría cuántica de campos. En física, una transición de fase describe un cambio brusco en las propiedades de un sistema, como cuando el agua hierve y se convierte en vapor. Hace años que se sabe que, durante una transición de fase, algunos sistemas físicos muy distintos acaban

obedeciendo las mismas leyes, en cuyo caso decimos que pertenecen a la misma «clase de universalidad». Este es un concepto muy potente en física que permite entender mucho mejor las propiedades de un sistema. Ahora, podemos explorar si aquellos algoritmos que son equivalentes a teorías de campos experimentan transiciones de fase durante su proceso de entrenamiento, y, en tal caso, a qué clase de universalidad pertenecen.

Aprender sobre campos cuánticos

Además de usar la física para entender mejor la inteligencia artificial, nuestros resultados también abren la puerta a usar el aprendizaje automático para aplicaciones físicas; por ejemplo, para aproximar teorías de campos intrincadas mediante otras más sencillas. Hoy por hoy, las aplicaciones del aprendizaje automático a la teoría de campos pueden considerarse aún en sus inicios, por lo que esta equivalencia podría inspirar líneas de investigación totalmente nuevas.

Por último, esta conexión entre la física fundamental y el aprendizaje automático puede relacionarse con una cuestión muy profunda y hasta ahora no resuelta: ¿cuáles son los fundamentos matemáticos de la teoría cuántica de campos?

En las últimas décadas, esta pregunta se ha explorado desde múltiples perspectivas. Una de ellas es la conocida como «teoría cuántica de campos constructiva», una parte de la cual estudia la cues-

tion desde el punto de vista de los campos aleatorios de Márkov. Por tanto, esta conexión entre la teoría de campos y el aprendizaje automático abre la oportunidad de iniciar un diálogo en la intersección de la física, la computación y las matemáticas; uno que podría acabar transformando las tres disciplinas.

Dimitrios Bachtis es físico teórico del Departamento de Matemáticas de la Universidad de Swansea.

PARA SABER MÁS

Construction of quantum fields from Markoff fields, Edward Nelson en *Journal of Functional Analysis*, vol. 12, págs. 97-112, enero de 1973.

Spin glass theory and beyond: An introduction to the replica method and its applications, Marc Mézard, Giorgio Parisi y Miguel Ángel Virasoro. World Scientific, 1987.

Quantum field-theoretic machine learning, Dimitrios Bachtis, Gert Aarts y Biagio Lucini en *Physical Review D*, vol. 103, art. 074510, abril de 2021.

EN NUESTRO ARCHIVO

Problemas físicos con muchas escalas de longitud, Kenneth G. Wilson en *lyC*, octubre de 1979.

Quarks por ordenador, Donald H. Weingarten en *lyC*, abril de 1996.

Las leyes del azar y el libre albedrío, Bartolo Luque en *lyC*, junio de 2021.

El misterio del muon, Lucius Bushnaq, Gregorio Herdoíza y Marina Krstić Marinković en *lyC*, junio de 2021.

INMUNOLOGÍA

El cerebro puede recordar y reactivar respuestas inmunitarias antiguas

El cerebro no solo ayuda a regular las reacciones inmunitarias, sino que también almacena y recupera «recuerdos» de ellas

ESTHER LANDHUIS

Los perros que oyen repetidamente una campana a la hora de comer generan una respuesta condicionada y salivan al escuchar el sonido de un simple tañido, según demostró el fisiólogo Iván Pávlov en la década de 1890. El cerebro de los animales aprende a asociar la campana con la comida y ordena a las

glándulas salivales a reaccionar en consecuencia.

Más de un siglo después, en un artículo publicado recientemente en *Cell*, la neuroinmunóloga Asya Rolls ha demostrado un tipo de condicionamiento similar en las respuestas del sistema inmunitario. Mediante técnicas genéticas

de última generación, ella y su equipo del Technion-Instituto Tecnológico de Israel, en Haifa, identificaron en ratones neuronas del cerebro que se activaban cuando les inducían experimentalmente una inflamación en el abdomen. A continuación demostraron que, cuando volvían a estimular esas neuronas, los