CECILIA RIKAP

TEORÍA DE LA DEPENDENCIA DIGITAL

Soberanía y desarrollo en el capitalismo del siglo XXI

CAJA NEGRA 03
FUTUROS PRÓXIMOS

Escaneado con CamScanner

INTELIGENCIA ARTIFICIAL: CÓDIGO, DATOS Y PROCESADORES

La IA como campo de investigación se ubica dentro de la matemática. Integra el grupo de modelos estadísticos que se utilizan para entender la estructura de un conjunto de datos y extraer de ellos patrones o regularidades. Los modelos de IA, entonces, son una técnica de medición y clasificación. Cada modelo está conformado por un conjunto de algoritmos, es decir, reglas o instrucciones para resolver un problema o realizar un cálculo. Usamos algoritmos para resolver problemas de lo más diversos. Uno de ellos es el de la big data: grandes volúmenes de datos para los cuales los métodos tradicionales de análisis estadístico resultan insuficientes o inútiles y que por eso han disparado la popularidad reciente de los algoritmos de IA. Una previsión (que seguramente se habrá quedado corta

Escaneado con CamScanner

porque en 2024 se crearon 173,4 zettabytes, equivalentes a un 41% más que en 2023) de la European Data Strategy estimaba que el volumen mundial de datos pasaría de 33 zettabytes en 2018 a 175 en 2025. 1 zettabyte equivale a 1.000.000.000.000.000.000.000 (10^{21}) bytes o a un trillón de gigabytes.[1] Basta compararlo con la capacidad de almacenamiento de nuestra computadora o celular. Por ejemplo, una Mac puede almacenar entre 128 y 8000 gigas dependiendo del modelo. Hablando de dispositivos, el 70% de la población mundial tiene conectividad móvil y hay más de 5300 millones de usuarios activos en internet, lo que equivale a alrededor del 68% de la población mundial.[2] Esto significa que aproximadamente el 70% de las personas está produciendo datos regularmente. Pero estos datos, recolectados libremente por plataformas y aplicaciones, son solo una parte del total de datos producidos a diario.

Las máquinas conectadas a internet, en inglés Internet of Things (internet de las cosas), y los satélites producen datos. Además, las organizaciones y personas producimos datos de forma cotidiana no solo como usuarios de internet: existen desde datos de producción, ventas y seguimiento de clientes de una empresa hasta datos de salud de una población. Todo el tiempo estamos generando datos y cada vez hay un mayor porcentaje de estos que se recopila y analiza para diversos fines.

Todos estos descomunales y crecientes volúmenes de datos son recolectados como datos crudos. Se utiliza este adjetivo para enfatizar que se trata de datos que aún no

1. Las cifras proyectas por la Comisión Europea están disponibles en la web de la European Data Strategy, en la sección "A Europe Fit for the Digital Age", disponible en ec.europa.eu.

2. Las cifras provienen de la Asociación Internacional de Telecomunicaciones y de un reporte de Cisco de 2020 que integra datos y proyecciones titulado "Cisco Annual Internet Report (2018-2023)", disponible en cisco.com.

Escaneado con CamScanner

fueron procesados, que no se transformaron en información o conocimiento. El acuerdo generalizado es que los datos solo permiten capturar valor cuando están agrupados –en bases estructuradas– y una vez procesados.

La técnica más avanzada para extraer conocimiento de datos hoy es usar un subconjunto de algoritmos de IA, una clase de *machine learning* llamada "aprendizaje profundo" o "redes neuronales profundas" (en inglés, *deep learning* o *deep neural networks*). Según el informe de la Organización Mundial de Propiedad Intelectual (WIPO, por sus siglas en inglés) de 2019,[3] esta es la tecnología de IA que más creció en solicitud de patentes entre 2012 y 2016 y su performance es mejor que la de cualquier otro modelo de IA.[4]

Estos modelos están constituidos por muchas capas que forman redes jerárquicas para aprender características y clasificar datos identificando patrones. Para ello, los algoritmos se entrenan procesando big data. El entrenamiento de un modelo de IA es como una suerte de puesta a punto. A diferencia de la IA que existía antes, los algoritmos de deep learning son modelos únicos y sumamente poderosos en materia económica porque tienen la capacidad de evolucionar. Se vuelven cada vez más expertos en función de la afluencia y el análisis de datos. En otras palabras, aprenden (de ahí el término *machine learning*) y mejoran por sí mismos a medida que procesan más datos, ajustando sus parámetros para arrojar previsiones más acertadas. A la hora de desarrollar y usar modelos de IA, el tamaño –de las bases de datos y del modelo– importa y mucho.

Así, datos y algoritmos han devenido complementarios estrictos; igualmente indispensables para que la IA

3. WIPO, "Artificial Intelligence", 2019, disponible en wipo.int.

4. Zahangir Alom y otros, "The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches", *ArXiv*, Cornell University, 2018, disponible en arxiv.org.

Escaneado con CamScanner

funcione. En general, cuanto mayor y más diversa sea la base de big data con la cual se entrena a un modelo de IA, mejor será la predicción. En la medida en que estos modelos se vuelven mejores con el uso, podemos pensarlos como medios de producción que se aprecian o valorizan con el tiempo. Mientras tanto, el resto de los medios de producción (máquinas y herramientas de todo tipo) se deprecian a medida que los usamos.

Basta pensar en los motores de búsqueda como Google. A lo largo de los años, las búsquedas fueron ganando precisión gracias al trabajo de retroalimentación del modelo de IA que está por detrás y que procesa más de 14 mil millones de búsquedas (nuevos datos) diarios. Cada búsqueda aumenta la principal fuente de ganancias de Google. Lo mismo sucede con los algoritmos de IA que se usan para fijar precios en las plataformas de comercio electrónico como Amazon y Mercado Libre. Para eso, se usan algoritmos de IA que ajustan continuamente los precios teniendo en cuenta datos de lo más variados. Estos algoritmos no solo fijan los precios de forma automática tomando en cuenta datos de ventas pasadas, sino que todo queda registrado: desde cómo movemos el mouse hasta qué cargamos en el carrito de compras, incluso si luego no concretamos la compra. El mismo modelo también tiene en cuenta la disponibilidad de los productos en los centros de distribución, las variables generales de la economía de cada país y los registros de datos pasados que las compañías ofrecen en sus plataformas para predecir la velocidad y efectividad con la cual se puede satisfacer la demanda. Nuestros datos de compras además nutren a algoritmos de recomendación que crean así perfiles de consumidores y nos inundan con publicidades personalizadas para que sigamos comprando.

Es tan variada la aplicación de estos modelos de deep learning que se ha comenzado a hablar de la IA como una tecnología de propósito general. Pero en realidad es algo

Escaneado con CamScanner

más: la IA es un nuevo método para inventar y crear.[5] Mediante el procesamiento de datos con IA es posible resolver preguntas de investigación muy diversas, realizar producciones artísticas y hasta fabricar hechos. Desde el descubrimiento de nuevos materiales o el mapeo de proteínas hasta las predicciones económicas, cada vez está más de moda usar algoritmos de deep learning para procesar grandes volúmenes de datos y así producir conocimiento.

En última instancia, el objetivo de las empresas que están detrás de estos modelos es producir una tecnología que no necesite procesos de adaptación, que pueda ser adoptada directamente en los entornos más variados. Ergo, que sus clientes potencialmente sean todas las organizaciones y personas del mundo. En esa búsqueda, apareció un tipo específico de modelo de deep learning que utiliza una técnica que se llama IA generativa. Se la denomina así porque cuando responde a estímulos (preguntas) del usuario genera contenidos (textos, imágenes, videos). El modelo interactúa con su entorno y mejora sus respuestas a partir de esas interacciones.

El ejemplo más conocido son los llamados grandes modelos de lenguaje, en inglés, *Large Language Models* (LLM). ChatGPT está basado en uno de los LLM más grande del mundo. Los LLM predicen qué palabra viene a continuación en una secuencia. Su utilización masiva supone que usuarios del mundo entero estén mejorando su funcionamiento de forma gratuita. La ambición por entrenar modelos LLM cada vez más grandes ha llevado a una

5. Stefano Bianchini y otros, "Artificial Intelligence in Science: An Emerging General Method of Invention", *Research Policy*, vol. 51, n° 10, 2022; Iain M. Cockburn y otros, *The Impact of Artificial Intelligence on Innovation*, National Bureau of Economic Research, 2018; Avi Goldfarb y otros, "Could Machine Learning Be a General Purpose Technology? A Comparison of Emerging Technologies Using Data from Online Job Postings", *Research Policy*, vol. 52, n° 1, 2023; Cecilia Rikap y Bengt-Åke Lundvall, *The Digital Innovation Race: Conceptualizing the Emerging New World Order*, Londres, Palgrave Macmillan, 2021.

Escaneado con CamScanner

utilización indiscriminada de datos de todo tipo, desde la colección de entradas de Wikipedia a los artículos y libros publicados por una de las editoriales más grandes del mundo, que fueron vendidos a Microsoft.[6]

El discurso dominante que se impone desde Silicon Valley es que los modelos de IA generativa pueden no solo responder sino incluso identificar preguntas de investigación y diseñar experimentos enteros. La promesa es que la IA podría automatizar el proceso mismo de descubrimiento reemplazando al trabajo creativo de investigación y desarrollo.

El procesamiento de datos con algoritmos de deep learning, sean de IA generativa o no, está muy por encima de las capacidades humanas de procesamiento de información. Sin embargo, no deja de ser matemática predictiva. En su forma más abstracta, aunque la complejidad haya cambiado, ambas hacen lo mismo. La modelización de la IA sigue consistiendo en construir una función matemática paramétrica que describa los datos de la mejor manera posible, de modo que pueda utilizarse para proporcionar inferencias o predicciones. La lógica es similar a cuando escribimos un correo electrónico, y un algoritmo de IA intenta predecir las siguientes palabras que usaremos en una frase. Los algoritmos de IA no razonan ni planifican, tampoco tienen emociones y, por lo tanto, no deciden considerando nada de lo anterior. Solo son –muy buenos– predictores.

El entrenamiento de modelos de IA de todo tipo y el almacenamiento y armado de las bases de big data usadas para ello requiere infraestructura digital. Hacen falta tanto servidores para almacenar al modelo y los datos como también procesadores –es decir, semiconductores– para

6. Se trata del catálogo de Routledge Books, parte del grupo Taylor & Francis. Ver Suswati Basu, "Academic Authors Alarmed over Taylor & Francis and Microsoft AI Deal", *How to Be*, 22 de julio de 2024, disponible en howtobe247.com.

Escaneado con CamScanner

seguir corriendo el modelo. Y una vez que el modelo está entrenado, su uso continuará necesitando todo un andamiaje de infraestructura digital. Esto incluye cables y satélites para transportar los datos de las interacciones de cada usuario con el modelo (por ejemplo, las preguntas que le hacemos a ChatGPT o a cualquier buscador). La misma infraestructura, a su vez, transporta sus respuestas en sentido inverso.

Los centros de datos equipados con un tipo de procesador llamado Unidad de Procesamiento Gráfico (GPU, por sus siglas en inglés) son el ámbito de entrenamiento por excelencia de los modelos de IA. Antes de probarse su potencia para entrenar estos modelos, las GPU se usaban principalmente para jugar videojuegos. En 2012, en NeurIPS, la conferencia académica de IA más famosa, académicos de la Universidad de Toronto presentaron a AlexNet, un modelo de reconocimiento de imágenes que había sido entrenado con GPU interconectadas. Aunque no era el primer modelo de machine learning entrenado de esta forma, recién entonces quedó claro que invertir en GPU para procesar más y más datos con modelos equipados con más parámetros resultaba en IA con mejor capacidad de predicción.

Desde entonces, el paradigma dominante entiende que cuanto mayor es la base de datos para procesar, más complejo el problema para resolver, por lo que el modelo de IA tendrá que incluir más parámetros y se necesitará más potencia de procesamiento para entrenarlo desde cero y ejecutarlo. Los modelos recientes que ahorran en procesamiento –como DeepSeek R1, o Gemma de Google– lo hacen mediante una técnica llamada destilado, que consiste en partir de modelos existentes para extraer sus aprendizajes. En concreto, usan menos procesadores porque no empiezan desde cero a entrenar el modelo, sino a partir de los aprendizajes de modelos existentes. Pero incluso en estos casos, todavía es necesario recurrir a GPU tanto para completar el entrenamiento como para correr el modelo.

Escaneado con CamScanner

Si hiciéramos una analogía con la preparación de una torta, producir IA requiere combinar en cantidades muy precisas tres ingredientes que son resultado del trabajo humano: algoritmos de estadística avanzada, datos e infraestructura digital (en particular, capacidad de procesamiento). Si falta uno, no hay torta. Y ninguno de ellos ni la IA resultante se parecen a las distopías de robots que conquistan el mundo y que es mejor dejar para la saga *Black Mirror*.

Escaneado con CamScanner