

WSJ

CONTENIDO LICENCIADO POR
 THE WALL STREET JOURNAL

RAFFAELE HUANG
 Wall Street Journal

Una compañía china de inteligencia artificial tiene a Silicon Valley maravillado por cómo sus programadores casi igualan a sus rivales estadounidenses, a pesar de usar chips inferiores. Los modelos de inteligencia artificial de DeepSeek, la empresa china, han ascendido al top 10 mundial en rendimiento, según un ranking popular, lo que sugiere que las restricciones a las exportaciones de Washington están teniendo dificultades para bloquear los rápidos avances de China.

El 20 de enero, DeepSeek presentó RL, un modelo especializado diseñado para la resolución de problemas complejos.

"DeepSeek RL es uno de los avances más sorprendentes e impresionantes que he visto", dijo Marc Andreessen, el capitán de riesgo de Silicon Valley que ha estado asesorando al Presidente Trump, en una publicación en X el viernes.

Los avances de DeepSeek provocaron una ola de ventas liderada por las acciones de chips el lunes por la mañana, ante las preocupaciones sobre si el enorme gasto de los gigantes tecnológicos estadounidenses en semiconductores de vanguardia y otras infraestructuras de inteligencia artificial estaba justificado. Los futuros vinculados al Nasdaq-100, centrado en la tecnología, cayeron un 4%, y Nvidia descendió más del 10% antes de la apertura del mercado.

El desarrollo de DeepSeek estuvo a cargo de un gestor de fondos de cobertura chino, Liang Wenfeng, que se ha convertido en el rostro del impulso de la inteligencia artificial en el país. El 20 de enero, Liang se reunió con el primer ministro de China y debatió sobre cómo las empresas locales podrían reducir la brecha con Estados Unidos.

Los especialistas afirman que la tecnología de DeepSeek aún está por detrás de la de OpenAI y Google, pero es un rival cercano a pesar de utilizar menos chips y menos avanzados, y en algunos casos saltándose pasos que los desarrolladores estadounidenses consideraban esenciales.

DeepSeek dijo que estrenar uno de sus últimos modelos costó US\$ 5,6 millones, en comparación con el rango de US\$ 100

Los modelos de inteligencia artificial de DeepSeek:

Silicon Valley está entusiasmado con un modelo de inteligencia artificial "hecho en China"

DeepSeek es calificado como "asombroso e impresionante" a pesar de funcionar con chips menos avanzados.

millones a US\$ 1.000 millones citados el año pasado por Dario Amodei, director ejecutivo del desarrollador de IA Anthropic, como el costo de construir un modelo.

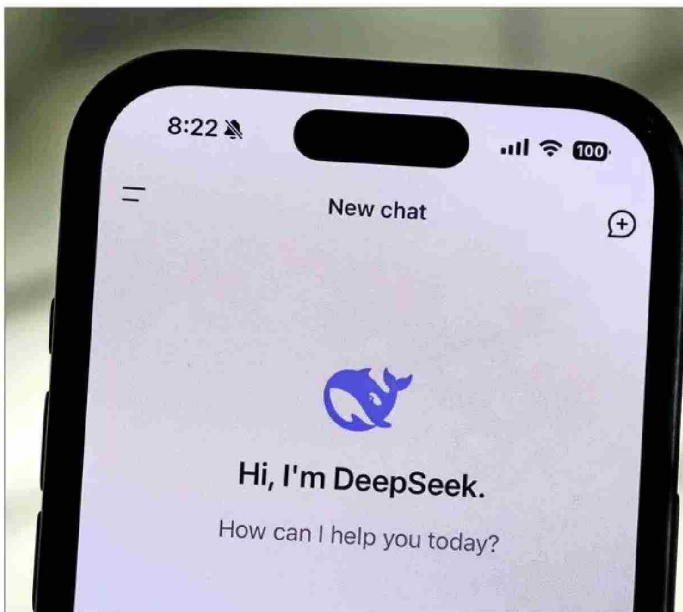
Barret Woodside, cofundador de la empresa de hardware de inteligencia artificial Positron de San Francisco, dijo que él y sus colegas han estado entusiasmados con DeepSeek. "Es genial", dijo Woodside, señalando los modelos de código abierto de DeepSeek en los que el código de software detrás del modelo de inteligencia artificial está disponible de forma gratuita.

Los usuarios del último modelo estrella de DeepSeek, llamado V3 y lanzado en diciembre, han notado que se niega a responder preguntas políticas delicadas sobre China y su líder Xi Jinping. En algunos casos, el producto ofrece respuestas en línea con la propaganda oficial de Beijing en lugar de incluir la perspectiva de los críticos del gobierno, como lo hace ChatGPT.

"El único problema es una censura poco elaborada por la República Popular China", dijo Woodside, pero agregó que esto podría eliminarse porque otros desarrolladores pueden modificar libremente el código.

DeepSeek dijo que tanto RL como V3 tuvieron un desempeño mejor o cercano a los principales modelos occidentales. Hasta el sábado, los dos modelos estaban clasificados entre los 10 primeros en Chatbot Arena, una plataforma organizada por investigadores de la Universidad de California, Berkeley, que califica el desempeño de los chatbots. Un modelo de Google Gemini estaba en el primer lugar, mientras que DeepSeek superó a Claude de Anthropic y a Grok de xAI de Elon Musk.

DeepSeek surgió de la unidad de investigación de IA de High-Flyer, un administrador de fondos de cobertura con US\$ 8 millones en activos que es cono-



Los especialistas afirman que la tecnología de DeepSeek aún está por detrás de la de OpenAI y Google, pero es un rival cercano, a pesar de utilizar menos chips y menos avanzados.

cido por aprovechar la IA para comercializar.

"Cuando los humanos toman decisiones de inversión, es un arte y lo hacen a su aire. Cuando los programas informáticos toman esas decisiones, es una ciencia y tienen la solución óptima", dijo Liang en un discurso de 2019.

Liang nació en 1985 y creció en la provincia de Guangdong, en el sureste de China. Asistió a la prestigiosa Universidad de Zhejiang y se especializó en visión artificial. Unos años después de graduarse, Liang fundó High-Flyer con dos amigos de la universidad en 2015.

Según personas cercanas a él, Liang prefiere que lo consideren ingeniero en lugar de comerciante. Su High-Flyer fue pionero en China en la aplicación del aprendizaje profundo al comercio computarizado. La técnica, basada en el cerebro humano, permite a las

computadoras analizar tipos de datos más diversos.

Un ejemplo sería una empresa que desea aprovechar la tecnología para dar respuestas de inteligencia artificial a las consultas de los clientes.

A principios del año pasado, DeepSeek redujo los precios de este servicio a una fracción de lo que cobraban otros proveedores, lo que llevó a la industria en China a iniciar una batalla de precios.

Anthony Poo, cofundador de una startup con sede en Silicon Valley que utiliza inteligencia artificial generativa para predecir los retornos financieros, dijo que su compañía pasó del modelo Claude de Anthropic a DeepSeek en septiembre. Las pruebas mostraron que DeepSeek tuvo un rendimiento similar por aproximadamente una cuarta parte del costo.

"El modelo de OpenAI es el

mejor en rendimiento, pero tampoco queremos pagar por capacidades que no necesitamos", afirmó Poo.

En su reunión del 20 de enero, Liang, de DeepSeek, le dijo al primer ministro chino, Li Qiang, que si bien las compañías chinas estaban trabajando para ponerse al día, las restricciones estadounidenses a la exportación de chips avanzados a China todavía eran un cuello de botella, según personas familiarizadas con la reunión.

En 2019, High-Flyer comenzó a construir un conjunto de chips para la investigación de la IA, en parte con fondos generados por su negocio de finanzas. La empresa ha dicho que más tarde construyó un conjunto más grande de alrededor de 10.000 unidades de procesamiento gráfico de Nvidia que se pueden utilizar para entrenar modelos de lenguaje de gran tamaño.

Solo un puñado de empresas en China tenían una infraestructura informática lo suficientemente potente como para desarrollar tales modelos a fines de 2022, cuando OpenAI lanzó ChatGPT.

DeepSeek afirmó en un informe técnico que utilizó un conjunto de más de 2.000 chips Nvidia para entrenar su modelo V3, en comparación con las decenas de miles de chips que se utilizan para entrenar modelos de tamaño similar. Algunos especialistas estadounidenses en inteligencia artificial han cuestionado recientemente si High-Flyer y DeepSeek están accediendo a una potencia de cálculo superior a la anunciada.

Algunos investigadores externos dijeron que el modelo DeepSeek carece de ciertas capacidades de sus rivales entrenados con mayor costo, por ejemplo, para rastrear el contexto de conversaciones largas.

Para su último modelo de razonamiento publicado el 20 de enero, DeepSeek se saltó un proceso conocido como ajuste fino supervisado, en el que los programadores incorporan el conocimiento de expertos humanos para darle al modelo una ventaja inicial. DeepSeek dijo que su modelo, diseñado para resolver problemas matemáticos complejos y desafíos similares, era comparable al modelo de razonamiento de OpenAI, aunque omitió el ajuste fino supervisado y se centró en el aprendizaje de refuerzo, es decir, en el ensayo y error dirigidos.

Jim Fan, un científico investigador de alto nivel de Nvidia, elogió como un gran avance el artículo de DeepSeek que informa sobre los resultados. Dijo en X que le recordaba a programas de inteligencia artificial pioneros anteriores que dominaban juegos de mesa como el ajedrez desde cero, sin imitar primero a grandes maestros.

Zack Kass, ejecutivo de OpenAI, dijo que los avances de DeepSeek, a pesar de las restricciones estadounidenses, "subrayan una lección más amplia: las limitaciones de recursos a menudo impulsan la creatividad".

Este artículo fue actualizado el lunes por la mañana. Se publicó originalmente el 25 de enero.

Traducido del inglés por "El Mercurio"

