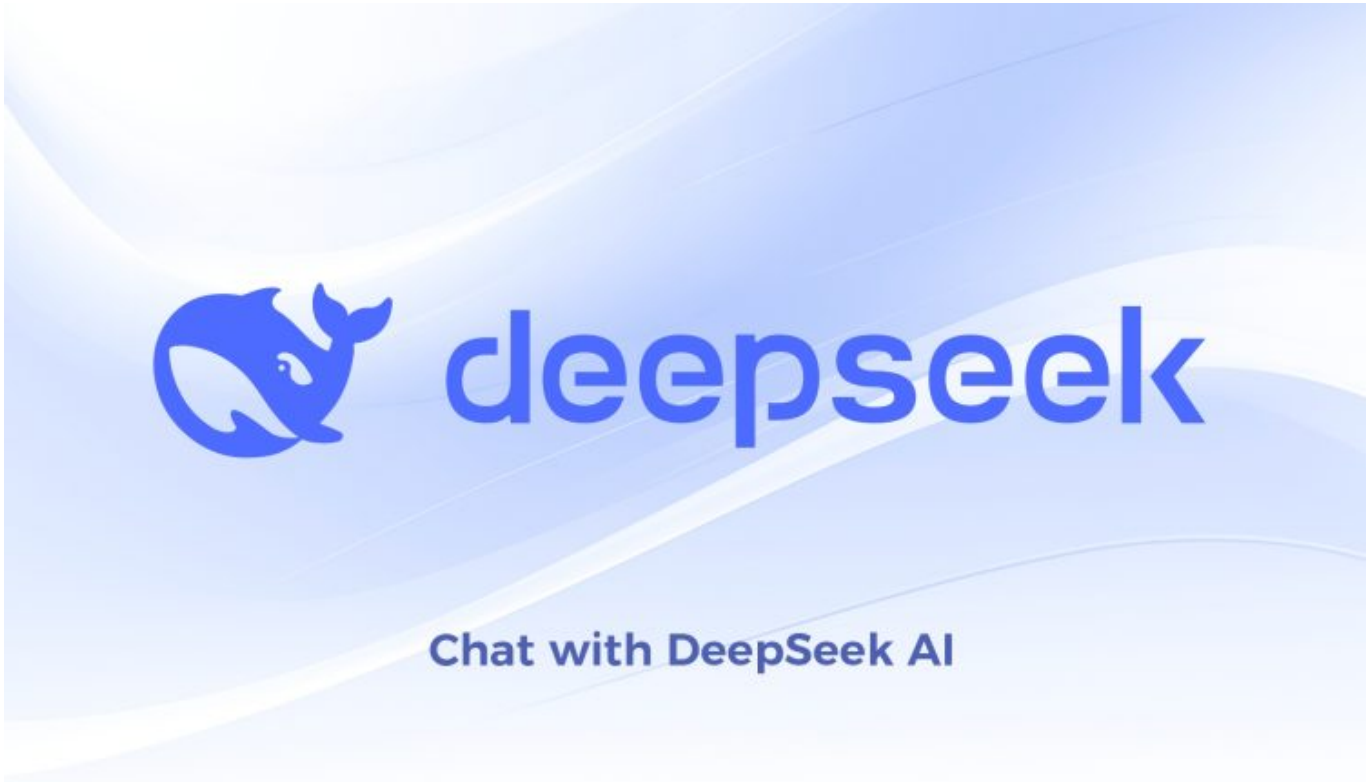
DeepSeek lanza un modelo de IA más eficiente y asequible que los de Meta y OpenAI

El modelo V3 de DeepSeek fue entrenado durante dos meses por 5.58 millones de dólares, por lo que utilizó significativamente menos recursos informáticos que sus rivales estadounidenses.

La *startup* china DeepSeek lanzó un nuevo **modelo de lenguaje grande** (LLM, por sus siglas en inglés), que causó revuelo en la industria global de la Inteligencia Artificial (IA), ya que las pruebas comparativas revelaron que **superó a los modelos rivales** de empresas como Meta Platforms, Llama, y el creador de ChatGPT, OpenAI.



La compañía con sede en Hangzhou dijo en una publicación de WeChat que la tercera versión de su LLM homónimo, **DeepSeek V3**, viene con 671 mil millones de parámetros y se entrena

alrededor de dos meses a un costo de 5.58 millones de dólares, utilizando significativamente menos recursos informáticos que los modelos desarrollados por empresas tecnológicas más grandes.

LLM se refiere a la tecnología que sustenta los servicios de IA Generativa como ChatGPT. En Inteligencia Artificial, una gran cantidad de parámetros es fundamental para permitir que un modelo se adapte a patrones de datos más complejos y haga predicciones precisas.

El desarrollo de un potente LLM por parte de DeepSeek con una fracción del desembolso de capital que suelen invertir empresas más grandes como Meta y OpenAI, muestra hasta qué punto han progresado las empresas de IA chinas, a pesar de las sanciones estadounidenses que han bloqueado su acceso a los semiconductores avanzados utilizados para entrenar modelos.

Aprovechando la nueva arquitectura diseñada para lograr un entrenamiento rentable, DeepSeek necesitó sólo 2.78 millones de horas de GPU, la cantidad total de tiempo que utiliza una unidad de procesamiento gráfico para entrenar un LLM, para su modelo V3. El proceso de entrenamiento de la empresa emergente utilizó las GPU H800 de Nvidia diseñadas para China.

Este proceso fue sustancialmente menor que los 30.8 millones de horas de GPU que la empresa matriz de Facebook, Meta, necesitó para entrenar su modelo Llama 3.1 en los chips H100 más avanzados de Nvidia, cuya exportación a China no está permitida.

Fuente: dplnews.