

## Impatto della Notizia su DeepSeek: Confronto con Meta e Implicazioni per AI & Semiconduttori

DeepSeek (DS) è interamente di proprietà di High-Flyer, un quant fund cinese (sembra che gestisca 8 bn\$) specializzato in AI con un track record di grande successo fondato da Liang WenFeng. High-Flyer ha fondato DeepSeek nell'aprile 2023 con l'obiettivo di concentrarsi su AGI (Artificial General Intelligence) e LLM (Large Language Models).

- DeepSeek V2, lanciato a maggio 2024, ha registrato un costo di soli 2 RMB per token generato, un valore estremamente competitivo rispetto agli attuali standard di mercato.
- Il modello ha ottenuto il settimo posto nella classifica LLM dell'Università di Waterloo, un riconoscimento importante che ne conferma le capacità.
- DeepSeek V4, lanciato il mese scorso, è stato addestrato su un dataset di 14,8 trilioni di token (contro i 13 trilioni di GPT-4o), con un costo di training di appena 5,6 milioni di dollari, assumendo un costo di \$2 per ora di noleggio di GPU H800.

Questi numeri evidenziano l'incredibile efficienza in termini di costo e potenza di calcolo del modello, rendendolo un potenziale game changer nel panorama AI. DeepSeek è già presente sull'Apple Store e scaricabile gratuitamente (a).

L'impatto che vediamo sui mercati finanziari deriva dal potenziale cambiamento di paradigma nell'AI. Tentiamo di capire qual è il cambiamento di paradigma su cui il mercato inizia a interrogarsi.

DeepSeek ha sviluppato un modello open-source che, a parità di prestazioni con GPT-4o, utilizza una quantità significativamente inferiore di potenza computazionale e GPU. Questo mette in discussione il modello di business attuale dell'AI, che si basa su ingenti investimenti in chip avanzati e infrastrutture a elevato consumo energetico.

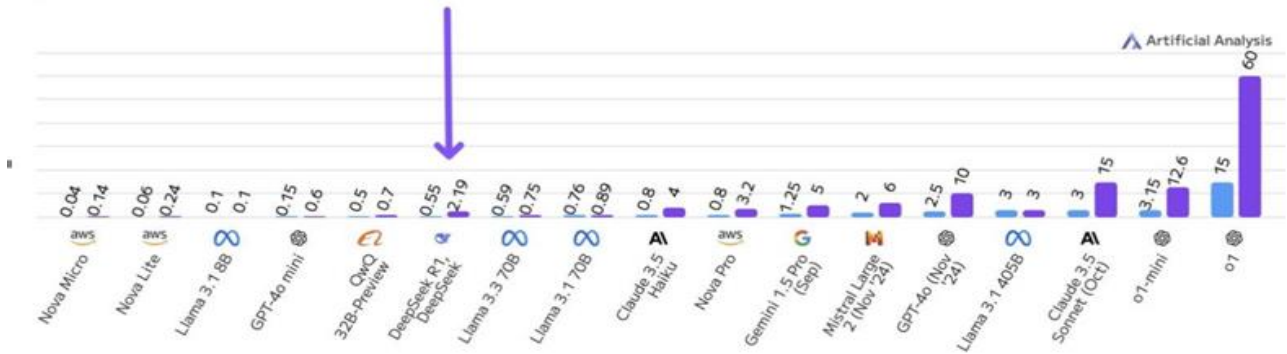
Confronto con Meta e il Vantaggio Competitivo di DeepSeek

- Efficienza nei costi:
  - DeepSeek-R1 offre un costo di \$0.14 per milione di token in input (cache hit), \$0.55 per cache miss e \$2.19 per output, rendendolo 13-107 volte più economico rispetto ai modelli OpenAI (GPT-4o).

## Pricing: Input and Output Prices

USD per 1M Tokens

■ Input price ■ Output price

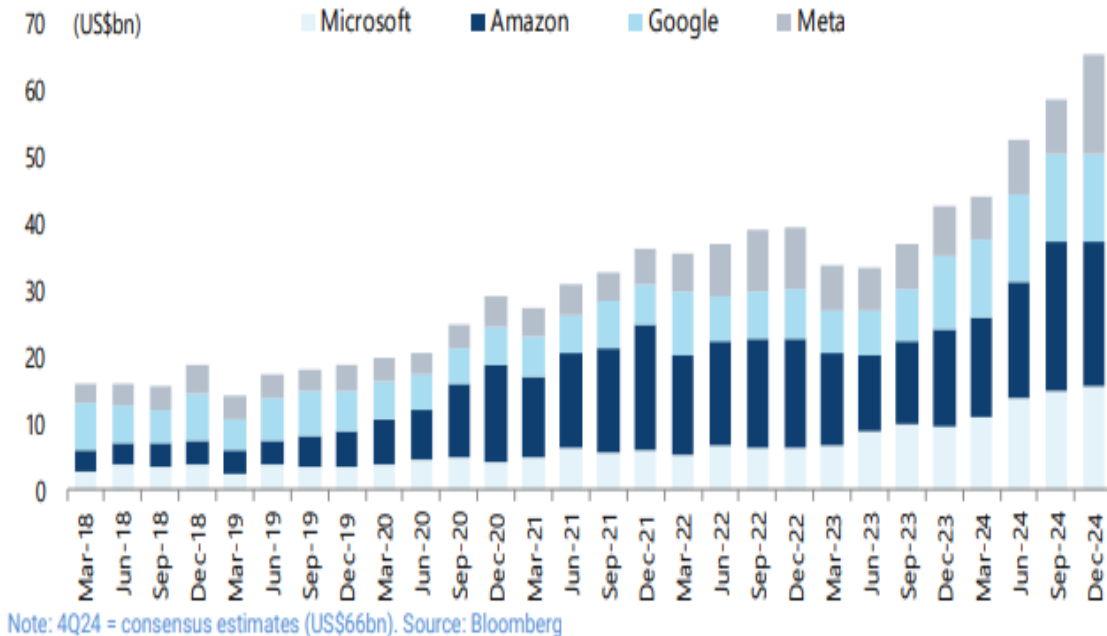


- In confronto, Meta ha speso oltre il 90% in più per addestrare il suo modello Llama 3.1, con costi di training stimati molto superiori.
- Ottimizzazione computazionale:
  - DeepSeek utilizza Mixture of Experts (MoE) e Multi-head Latent Attention (MLA), attivando solo 20 miliardi di parametri per query su un totale di 200 miliardi, riducendo drasticamente i costi di inferenza e i tempi di risposta.
  - Meta e altri big tech AI come OpenAI e Google, al contrario, si affidano ancora a modelli monolitici con un uso intensivo di GPU.
- Costi di training:
  - Il costo di addestramento di DeepSeek V4 è stato di soli \$5.6 milioni, contro le centinaia di milioni (se non miliardi) spesi da aziende come Meta, OpenAI e Google per modelli simili.
  - Questo rappresenta meno del 10% del costo di addestramento di Meta Llama 3.1.

### Possibili implicazioni sul Mercato:

- Pressione sulle azioni dei semiconduttori e dei data center:
  - La minore necessità di potenza di calcolo potrebbe ridurre la domanda di GPU avanzate (Nvidia), di foundries dei chips avanzati (TSMC) e delle società operanti nelle infrastrutture data center.
- Ripercussioni sulla supply chain:
  - Aziende giapponesi che forniscono componenti a Nvidia sono già in calo, così come il settore del liquid cooling in Cina, mentre il settore internet in Hong Kong è in rialzo.

- Focus sull'efficienza dell'AI e sostenibilità degli investimenti:
  - DeepSeek potrebbe segnare un punto di svolta nel settore, costringendo il mercato a riconsiderare l'enorme Capex in GPU (vedere chart sulle stime capex dei main hyperscalers US).



- Se il trend dovesse consolidarsi, le previsioni di spesa AI per il 2026 potrebbero subire un drastico ridimensionamento.
- Implicazioni geopolitiche:
  - La Cina ha sviluppato DeepSeek con risorse limitate a causa delle restrizioni USA sui chip avanzati. Questo potrebbe accelerare l'innovazione cinese e creare ulteriore pressione sugli Stati Uniti nel definire future politiche di esportazione dei semiconduttori.

## Conclusione

DeepSeek sta emergendo come un potenziale disruptor nel settore AI, con il rischio di un ripensamento del modello di investimento attuale. Il mercato potrebbe passare da una strategia di "più potenza computazionale a qualsiasi costo" a un focus sull'efficienza e ROI, riducendo la domanda di chip avanzati e data center.

Sebbene sia ancora presto per valutarne l'impatto definitivo, la notizia ha già generato un'ondata di vendite nei titoli dei semiconduttori e data center. Se questo trend si confermerà, potrà portare a una profonda revisione delle strategie di investimento nel settore AI che si tradurrà nel rimettere in discussione le stime sui ritorni oggi attesi e sulle valutazioni che si basano proprio su quei ritorni attesi.

## DISCLAIMER

Le presenti informazioni sono state elaborate da Valori Asset Management SA (Svizzera), autorizzata e regolamentata dall'Autorità federale di vigilanza sui mercati finanziari FINMA.

Questo materiale viene fornito da Valori Asset Management SA (Svizzera) esclusivamente a scopo informativo, è destinato esclusivamente al vostro uso e non costituisce un'offerta o un impegno, una sollecitazione di un'offerta o di un impegno, né un consiglio o una raccomandazione a effettuare o concludere alcuna transazione (alle condizioni indicative indicate o in altro modo).

Questo materiale è stato elaborato da Valori Asset Management SA (Svizzera) sulla base di ipotesi e parametri determinati in buona fede. Le ipotesi e i parametri utilizzati non sono gli unici che potrebbero essere utilizzati e pertanto non si garantisce l'accuratezza, la completezza o la ragionevolezza di tali quotazioni, divulgazioni o analisi. Una serie di ulteriori ipotesi o parametri, o altri fattori di mercato e altre considerazioni, potrebbero portare a differenti analisi o valutazioni in buona fede delle transazioni sopra descritte. I risultati passati non devono essere considerati come un'indicazione o una garanzia dei risultati futuri e non viene fornita alcuna dichiarazione o garanzia, esplicita o implicita, in merito ai risultati futuri. Le opinioni e le stime possono essere modificate senza preavviso. Le informazioni sopra riportate sono state ottenute o basate su fonti ritenute affidabili da Valori Asset Management SA (Svizzera), ma Valori Asset Management SA (Svizzera) non ne dichiara né garantisce l'accuratezza o la completezza. Questo materiale non ha la pretesa di contenere tutte le informazioni che un soggetto interessato potrebbe desiderare. In ogni caso, le parti interessate dovrebbero condurre le proprie indagini e analisi delle transazioni descritte in questo materiale e dei dati in esso contenuti.

Valori Asset Management SA (Svizzera) può, di volta in volta, partecipare o investire in altre operazioni di finanziamento con gli emittenti dei titoli citati nel presente documento, prestare servizi per tali emittenti o sollecitarne l'attività, e/o avere una posizione o effettuare operazioni sui titoli o sui loro derivati.