

CNN, 2025. 1. 28

中 DeepSeek, 초고효율 AI 모델 출시로 대규모 AI 투자 패러다임에 변화



중국 AI 개발사 DeepSeek가 기존의 1/10이 채 안되는 개발 비용으로

GPT-4 수준의 고성능 LLM 모델인 'DeepSeek-R1' 출시



MoE, IBRL 등의 기술 응용으로, 저비용·고성능의 모델을 구현한 것이 특징이며,

DeepSeek는 기존 AI 산업의 경쟁 패러다임 변화를 예고

▶ DeepSeek, 초고효율·저비용의 고성능 오픈소스 기반 AI 모델 공개

DeepSeek의 거대 언어모델 'R1'은 약 6백만 달러 미만의 비용으로 개발된 저비용·고효율의 첫 GPT 경쟁 모델이다. DeepSeek-R1의 성능 뒤에는 전문가 조합 (MoE, Mixture of Experts) 아키텍처와 추론 기반 강화학습 (IBRL, Inference-Based Reinforcement Learning) 기반의 효율적 설계가 핵심으로 작용했다. 이를 통해 DeepSeek는 학습 효율을 높여 성능을 유지하면서도 자원 사용을 최소화하는 데 성공했다. 이러한 혁신으로 R1은 고성능 모델 개발에 막대한 자원이 반드시 필요하다는 기존 틈틈에 균열을 내며, 샘 올트먼, 마크 저커버그, 일론 머스크 등이 주장해온 '자원 중심 AI 개발 패러다임'을 흔들 수 있음을 시사했다.

▶ 혁신적인 비용 절감과 규모의 비경제성 입증으로 시장의 격변 예고

DeepSeek의 등장은 "수만 개의 칩과 초대형 데이터센터가 있어야 고성능 모델을 만들 수 있다"는 기존 AI 산업의 전제에 의문을 제기했다. NVIDIA 칩으로 데이터센터를 무한 확장하는 방식이 경제적으로 비효율적일 수 있음을 보여준 것이다. 그동안 미국 빅테크 기업들은 수십억~수백억 달러 규모의 GPU·전력·데이터 비용을 투입해 모델을 개발하고, 출시 후에도 막대한 추론 비용을 감당해야 했다. 하지만 DeepSeek는 적은 연산 자원과 에너지로도 고성능 모델 개발이 가능하다는 점을 입증했다. 이는 규모의 경제 기반의 기존 AI 산업 전략이 더 이상 절대적이지 않음을 보여주며, 경쟁 패러다임이 변화할 수 있음을 시사한다.

▶ AI 골드러시의 핵심 '삼과 곡괭이' 전략 재평가

AI 골드러시에서 '삼과 곡괭이'를 공급하는 기업, 즉 GPU와 인프라를 제공하는 기업들도 AI 서비스를 개발하는 기업만큼 산업의 중심에 서있다. NVIDIA는 그 대표 기업으로, 수년간 AI 산업에 필수적인 칩을 공급하며 시가총액 3조 달러 규모로 성장했다. 그러나 DeepSeek가 적은 자원으로 경쟁력 있는 모델을 구축하자, 칩 수요가 지금처럼 무한정 늘어나지 않을 수 있다는 우려가 제기됐다. NVIDIA는 DeepSeek 모델을 "AI 발전의 뛰어난 사례"라고 평가했지만, 월스트리트 일부 분석가들은 주가 급락세를 일시적인 반응으로 보고, 장기적으로는 칩 수요가 유지될 것이라고 전망한다.

1월의 용어 전문가 조합, 희소 어텐션, 인간 피드백 기반 강화학습, AI 피드백 기반 강화학습, 추론 기반 강화학습

출처 : 1) CNN(2025. 1. 28), DeepSeek just blew up the AI industry's narrative that it needs more money and power

2) Dev. to(2025. 1. 28.), DeepSeek and the Power of Mixture of Experts (MoE)