

DeepSeek-R1: Distillation, Training, and the Flawed Panic in the AI Industry

Abstract

The emergence of DeepSeek-R1 has sparked concerns within the AI industry, primarily due to claims of significantly lower training costs compared to competitors like OpenAI. This report delves into the concept of knowledge distillation in Large Language Models (LLMs), specifically how it was employed in training DeepSeek-R1. Furthermore, it analyzes the validity of the industry's cost-related panic, arguing that while DeepSeek has achieved notable efficiency gains, the narrative surrounding its disruptive potential is overblown and based on incomplete information. The report critically evaluates DeepSeek's training methodology, cost-saving strategies, and performance benchmarks, concluding that while the model represents a step forward in efficiency, it doesn't constitute a paradigm shift that threatens the established players in the AI landscape.

Introduction

The development of LLMs has been marked by a relentless pursuit of scale, with models growing increasingly larger and more computationally expensive. This trend has raised concerns about accessibility and sustainability, prompting research into model compression techniques like knowledge distillation. Knowledge distillation involves transferring the "knowledge" from a large, complex "teacher" model to a smaller, more efficient "student" model ([Knowledge distillation: a way to make a large model more efficient and accessible](#)). DeepSeek-R1 leverages this technique, alongside reinforcement learning, to achieve impressive performance while maintaining lower training costs. However, the narrative surrounding DeepSeek's cost efficiency and its potential to disrupt the AI industry warrants a critical examination.

Knowledge Distillation in LLMs

Knowledge distillation allows smaller models to inherit the complex behaviors and reasoning capabilities of larger models without the associated computational burden. Instead of simply mimicking the teacher model's outputs, distillation aims to replicate its underlying "thought processes" ([Knowledge distillation: a way to make a large model more efficient and accessible](#)). This is achieved by training the student model not just on the ground truth labels but also on the softer probability distributions outputted by the teacher model. This allows the student to learn from the teacher's nuanced understanding of the data, including its uncertainties and biases. In the context of LLMs, this translates to transferring stylistic nuances, reasoning abilities, and even alignment with human values ([Knowledge distillation: a way to make a large model more efficient and accessible](#)).

DeepSeek-R1's Training Methodology

DeepSeek-R1's training employed a multi-stage approach combining supervised fine-tuning (SFT) and reinforcement learning (RL) ([Bite: How Deepseek R1 was trained](#)). Initially, the base model, DeepSeek V3, underwent SFT using a dataset of chain-of-thought (CoT) data generated by both the R1-zero model and human annotators. This stage focused on improving readability and coherence. Subsequently, RL was applied, concentrating on reasoning-intensive tasks like coding and mathematics, utilizing rule-based reward models and an additional reward for language consistency ([Bite: How Deepseek R1 was trained](#)). This multi-stage process, alternating between SFT and RL, allowed DeepSeek-R1 to refine its reasoning capabilities while maintaining readability and alignment with human preferences ([DeepSeek R1: It's All About Architecture and Training Approach](#)). Distillation played a crucial role in creating smaller, more efficient versions of the R1 model without requiring the computationally expensive RL training ([DeepSeek-R1: Revolutionizing Reasoning with Reinforcement Learning and Distillation](#)).

The Flawed Panic: Deconstructing the Cost Narrative

While DeepSeek has publicized a significantly lower training cost for its models (e.g., \$6 million for V3), the narrative surrounding

this figure is misleading. Reports suggest DeepSeek leveraged a substantial stockpile of Nvidia chips (potentially costing around \$1 billion), which were not accounted for in their publicized figures due to U.S. export control restrictions ([Is the DeepSeek Panic Overblown?](#)). This omission significantly skews the cost comparison with competitors like OpenAI, who reportedly spent over \$100 million training GPT-4 ([Is the DeepSeek Panic Overblown?](#)).

Furthermore, DeepSeek's lower model access fees (\$2.19 per million tokens compared to OpenAI's \$60) don't necessarily reflect superior cost efficiency. Experts suggest this pricing strategy could be a loss-leader tactic to gain market share, potentially operating at a loss on inference ([Is the DeepSeek Panic Overblown?](#)). DeepSeek's cost-saving measures, such as proprietary energy-efficient accelerators and data optimization, contribute to their lower expenses ([DeepSeek Vs OpenAI: A comparative analysis of LLM development and cost efficiency](#)). However, these efficiencies, while noteworthy, do not represent a fundamental technological breakthrough that invalidates the investments made by other leading AI companies.

DeepSeek-R1's Performance: A Balanced Perspective

DeepSeek-R1 has demonstrated impressive performance on various benchmarks, including reasoning tasks (AIME 2024, MATH-500), general QA, and long-context understanding (AlpacaEval, ArenaHard) ([DeepSeek-R1: Revolutionizing Reasoning with Reinforcement Learning and Distillation](#)). However, experts caution against overinterpreting these results. While R1 showcases advancements in efficiency and specific task performance, it's not considered a groundbreaking leap in AI capabilities ([Is the DeepSeek Panic Overblown?](#)).

Conclusion

DeepSeek-R1's utilization of knowledge distillation and reinforcement learning represents a significant step forward in LLM training efficiency. However, the narrative surrounding its disruptive potential based on significantly lower training costs is flawed and incomplete. While DeepSeek has undoubtedly achieved

notable efficiency gains, the undisclosed hardware investments and potentially loss-making pricing strategy cast doubt on the true extent of their cost advantage. Furthermore, while R1 performs well on various benchmarks, it doesn't represent a paradigm shift in AI capabilities that threatens the established players in the industry. The current panic within the AI industry is therefore overblown and warrants a more nuanced understanding of DeepSeek's achievements and limitations.

References

Gore, N. (2025, January). DeepSeek Vs OpenAI: A comparative analysis of LLM development and cost efficiency. Medium. <https://medium.com/@nrgore1/deepseek-vs-openai-a-comparative-analysis-of-llm-development-and-cost-efficiency-a8534f32c9a8>

Maheshwarappa, A. (2025, January). DeepSeek-R1: Revolutionizing Reasoning with Reinforcement Learning and Distillation. Medium. <https://abhishek-maheshwarappa.medium.com/deepseek-r1-revolutionizing-reasoning-with-reinforcement-learning-and-distillation-24f9e1877627>

Odunola, J. (2023, November 15). Exploring Knowledge Distillation in Large Language Models. Medium. https://medium.com/@jenrola_odun/exploring-knowledge-distillation-in-large-language-models-9d9be2bff669

Schmid, P. (n.d.). Bite: How Deepseek R1 was trained. <https://www.philschmid.de/deepseek-r1>

Singh, D. (2025, February 1). DeepSeek-R1: Redefining Open-Source Reasoning in LLMs. Medium. <https://medium.com/@deepankar080892/deepseek-r1-redefining-open-source-reasoning-in-llms-89f09250afed>

TeqnoVerse. (2025, January). DeepSeek R1: It's All About Architecture and Training Approach. Medium. <https://teqnoverse.medium.com/deepseek-r1-its-all-about-architecture-and-training-approach-50af74c223b8>

TIME. (n.d.). Is the DeepSeek Panic Overblown?. <https://time.com/7211646/is-deepseek-panic-overblown/>

Toloka. (n.d.). Knowledge distillation: a way to make a large model more efficient and accessible. <https://toloka.ai/blog/knowledge-distillation/>