DeepSeek R1: What Sets It Apart?

Article · February 2025

CITATIONS
0

READS
45

1 author:

Aditya Kumar
Indian Institute of Technology Madras
3 PUBLICATIONS **0** CITATIONS

SEE PROFILE

DeepSeek R1: What Sets It Apart?

Introduction

- DeepSeek R1 is redefining the landscape of artificial intelligence with its innovative design, cost-effectiveness, and open-source accessibility. Developed by **DeepSeek**, this large language model (LLM) stands out for its advanced reasoning capabilities and efficient architecture, challenging the dominance of proprietary AI systems.
- Utilizing a sophisticated Mixture of Experts architecture with 671 billion parameters, it delivers exceptional performance while maintaining remarkable cost-effectiveness. The model stands out by activating only a fraction of its parameters during operation, enabling high-efficiency computational processing.
- With impressive benchmark results across mathematical reasoning, coding, and general knowledge domains, DeepSeek R1 offers developers and researchers an accessible, transparent, and powerful AI solution. Its open-source nature democratizes advanced AI technology, breaking down barriers and fostering collaborative innovation in the rapidly evolving artificial intelligence landscape.

This article delves into DeepSeek R1's origins and development, explores its innovative architecture, highlights why it is cost-effective, examines its performance benchmarks, and discusses the advantages its open-source nature brings to the AI community.

Origin, Development, and Its Architecture

Origin

DeepSeek R1 originated from DeepSeek, a Chinese AI startup founded in July 2023 by Liang Wenfeng, a graduate of Zhejiang University specializing in information and electronic engineering. The company is headquartered in Hangzhou, China, and was incubated by **High-Flyer**, a hedge fund that Liang had previously founded in 2015. [1] Similar to OpenAI's Sam Altman, Liang's goal is to develop artificial general intelligence (AGI). **R1- 0** is the first reasoning model of DeepSeek.

Development

DeepSeek released its first model in November 2023, but it wasn't until January 2025 that the company gained global attention with the release of R1.

Model	Release Date
DeepSeek Coder	November 2023
DeepSeek LLM	December 2023
DeepSeek-V2	May 2024
DeepSeek-Coder V2	July 2024
DeepSeek-V3	December 2024
DeepSeek-R1	January 2025
Janus-Pro-7B	January 2025

R1 builds upon its predecessor, DeepSeek-R1-Zero, which was trained entirely through reinforcement learning.

Reinforcement Learning

Reinforcement Learning (RL) is a branch of machine learning focused on making decisions to maximize cumulative rewards in a given situation. Unlike supervised learning, RL involves learning through experience. In RL, an agent learns to achieve a goal in an uncertain, potentially complex environment by performing actions and receiving feedback through rewards or penalties.

DeepSeek-R1 (Development technique)

Reinforcement Learning with Cold Start:

DeepSeek-R1-Zero struggles with challenges like poor readability, and language mixing. To make reasoning processes more readable and share them with the open community, DeepSeek explores a method that utilizes RL with a human-friendly cold start (Fine-tuning on carefully crafted Chain of Thought (CoT) reasoning examples to improve clarity and readability) in R1.

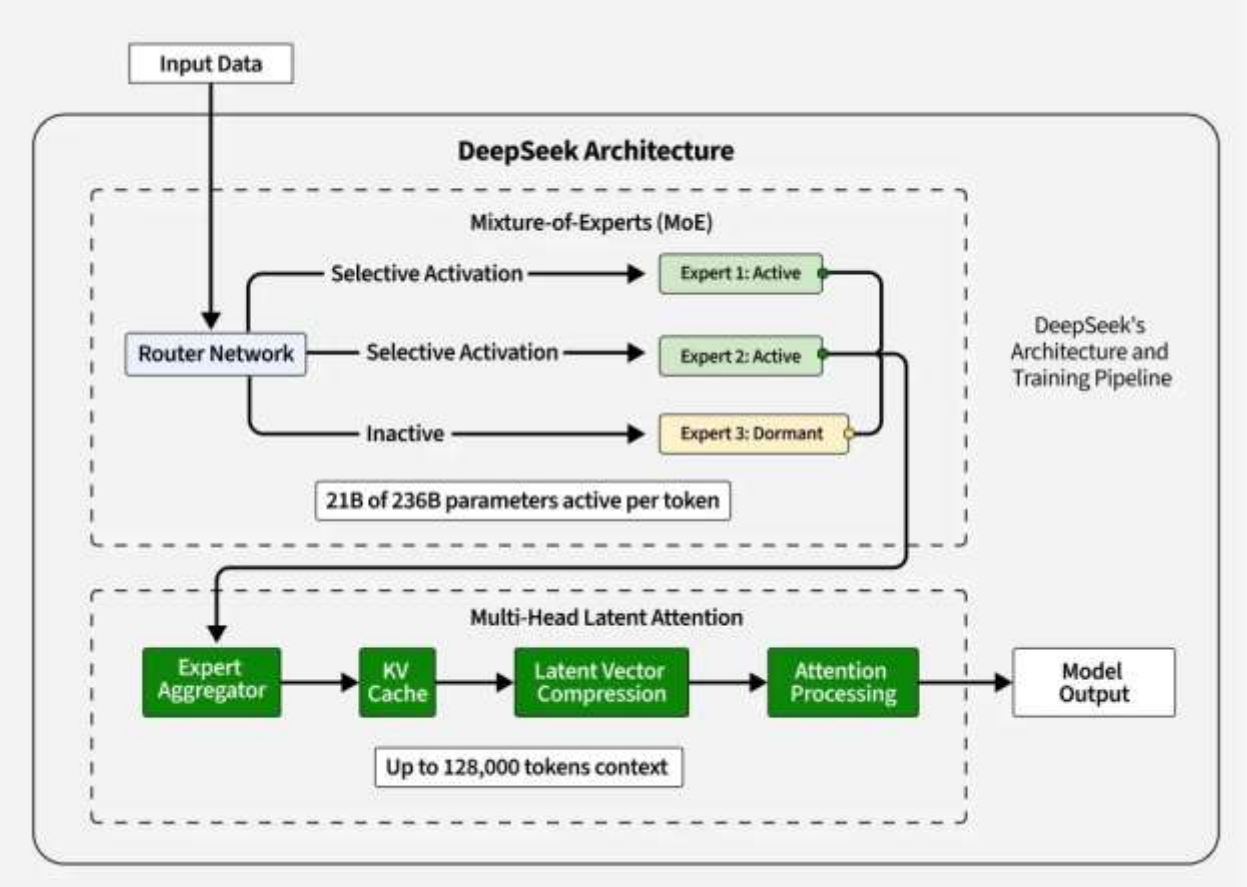
Iterative reinforcement learning:

A large-scale reinforcement learning training process is employed in DeepSeek R1, this phase focuses on enhancing the model's reasoning capabilities, particularly in reasoning-intensive tasks such as coding, mathematics, science, and logic reasoning. However, in the training process, CoT often exhibits language mixing which can be mitigated through rewarding during RL training.

Supervised Fine-Tuning:

DeepSeek R1 incorporates limited Supervised Fine-Tuning (SFT) in its training process. It's used for initial guidance and post-reinforcement learning refinement, enhancing readability and domain-specific performance while preserving RL-driven reasoning capabilities.

Architecture



- **Mixture of Experts (MoE) Architecture:**

DeepSeek R1 employs a Mixture of Experts (MoE) framework, which divides the model into specialized sub-models (experts). During inference, only a subset of these experts is activated based on the task, significantly reducing computational costs while maintaining high performance.

Out of its 671 billion parameters, only 37 billion are activated per forward pass, ensuring efficiency without compromising quality.

This sparsity-based approach allows the model to handle diverse tasks effectively and reduces inference-time computation costs.

- **Multi-Level Attention (MLA):**

The model integrates a Multi-Level Attention (MLA) mechanism that enhances its ability to process complex inputs by focusing on different levels of information simultaneously.

MLA is particularly useful for tasks requiring nuanced understanding, such as natural language processing and decision-making.

- **Transformer-Based Design:**

In addition to MoE, DeepSeek-R1 incorporates advanced transformer layers for natural language processing. These layers incorporate optimizations like sparse attention mechanisms and efficient tokenization to capture contextual relationships in text, enabling superior comprehension and response generation.

Why is it cost-effective?

- **Hardware Optimization**: R1 was reportedly trained on just a couple thousand H800 chips — a cheaper and less powerful version of Nvidia's $40,000 H100 GPU.
- **Efficient Architecture**: DeepSeek R1 uses a Mixture of Expert (MoE) architecture, activating only 37 billion out of 671 billion parameters per forward pass, significantly reducing computational costs.
- **Lower API Pricing**: DeepSeek's API costs are substantially lower than competitors. For 1 million tokens, DeepSeek charges $0.42, compared to $90 for GPT-3.5 Turbo and $180 for GPT-4 Turbo.

- **Group Relative Policy Optimization:**

Group Relative Policy Optimization In order to save the training costs of RL, we adopt Group Relative Policy Optimization (GRPO) (Shao et al., 2024), which foregoes the critic model that is typically the same size as the policy model, and estimates the baseline from group scores instead. Specifically, for each question q, GRPO samples a group of outputs $\{o_1, o_2, \cdots, o_G\}$ from the old policy $\pi_{\theta_{old}}$ and then optimizes the policy model π_θ by maximizing the following objective:

$$\mathcal{J}_{GRPO}(\theta) = \mathbb{E}[q \sim P(Q), \{o_i\}_{i=1}^{G} \sim \pi_{\theta_{old}}(O|q)]$$

$$\frac{1}{G}\sum_{i=1}^{G}\left(\min\left(\frac{\pi_\theta(o_i|q)}{\pi_{\theta_{old}}(o_i|q)}A_i, \text{clip}\left(\frac{\pi_\theta(o_i|q)}{\pi_{\theta_{old}}(o_i|q)}, 1-\varepsilon, 1+\varepsilon\right)A_i\right) - \beta\mathbb{D}_{KL}\left(\pi_\theta||\pi_{ref}\right)\right), \quad (1)$$

$$\mathbb{D}_{KL}\left(\pi_\theta||\pi_{ref}\right) = \frac{\pi_{ref}(o_i|q)}{\pi_\theta(o_i|q)} - \log\frac{\pi_{ref}(o_i|q)}{\pi_\theta(o_i|q)} - 1, \quad (2)$$

where ε and β are hyper-parameters, and A_i is the advantage, computed using a group of rewards $\{r_1, r_2, \ldots, r_G\}$ corresponding to the outputs within each group:

$$A_i = \frac{r_i - mean(\{r_1, r_2, \cdots, r_G\})}{std(\{r_1, r_2, \cdots, r_G\})}. \quad (3)$$

- **Open-Source Approach:** Being open-source allows users to customize and deploy the model on their own infrastructure, potentially eliminating recurring API costs.

Model	Input Cost (per 1M tokens)	Output Cost (per 1M tokens)	Performance Highlights
DeepSeek R1	$0.55 (Cache Miss)	$2.19	- AIME 2024: 79.8%
	$0.14 (Cache Hit)		- MATH-500: 97.3%
			- Codeforces: 96.3 percentile
OpenAI o1 Mini	$3 - $5	$12 - $15	Comparable to DeepSeek R1 on industry benchmarks
OpenAI o1	Not specified	$60	- AIME 2024: 79.2%
			- MATH-500: 96.4%
			- Codeforces: 96.6 percentile

Performance Benchmarks

DeepSeek conducted comprehensive evaluations against several strong baselines. For more information click here.

	Benchmark (Metric)	Claude-3.5-Sonnet-1022	GPT-4o 0513	DeepSeek V3	OpenAI o1-mini	OpenAI o1-1217	DeepSeek R1
	Architecture	-	-	MoE	-	-	MoE
	# Activated Params	-	-	37B	-	-	37B
	# Total Params	-	-	671B	-	-	671B
English	MMLU (Pass@1)	88.3	87.2	88.5	85.2	**91.8**	90.8
	MMLU-Redux (EM)	88.9	88.0	89.1	86.7	-	**92.9**
	MMLU-Pro (EM)	78.0	72.6	75.9	80.3	-	**84.0**
	DROP (3-shot F1)	88.3	83.7	91.6	83.9	90.2	**92.2**
	IF-Eval (Prompt Strict)	**86.5**	84.3	86.1	84.8	-	83.3
	GPQA Diamond (Pass@1)	65.0	49.9	59.1	60.0	**75.7**	71.5
	SimpleQA (Correct)	28.4	38.2	24.9	7.0	**47.0**	30.1
	FRAMES (Acc.)	72.5	80.5	73.3	76.9	-	**82.5**
	AlpacaEval2.0 (LC-winrate)	52.0	51.1	70.0	57.8	-	**87.6**
	ArenaHard (GPT-4-1106)	85.2	80.4	85.5	92.0	-	**92.3**
Code	LiveCodeBench (Pass@1-COT)	38.9	32.9	36.2	53.8	63.4	**65.9**
	Codeforces (Percentile)	20.3	23.6	58.7	93.4	**96.6**	96.3
	Codeforces (Rating)	717	759	1134	1820	**2061**	2029
	SWE Verified (Resolved)	**50.8**	38.8	42.0	41.6	48.9	49.2
	Aider-Polyglot (Acc.)	45.3	16.0	49.6	32.9	**61.7**	53.3
Math	AIME 2024 (Pass@1)	16.0	9.3	39.2	63.6	79.2	**79.8**
	MATH-500 (Pass@1)	78.3	74.6	90.2	90.0	96.4	**97.3**
	CNMO 2024 (Pass@1)	13.1	10.8	43.2	67.6	-	**78.8**
Chinese	CLUEWSC (EM)	85.4	87.9	90.9	89.9	-	**92.8**
	C-Eval (EM)	76.7	76.0	86.5	68.9	-	**91.8**
	C-SimpleQA (Correct)	55.4	58.7	**68.0**	40.3	-	63.7

Open-Source Advantage: What It Means for Developers and Researchers:

- R1 is fully open-source, allowing developers to freely examine, modify, and build upon the model.
- Lowering the Cost of Innovation: Being open-source, DeepSeek-R1 eliminates hefty licensing fees and paywalls that can choke a company's AI budget.
- This transparency fosters innovation and collaboration within the AI community.

- Open source DeepSeek-R1, as well as its API, will benefit the research community to distill better smaller models in the future.
- Enabling startups, academic institutions, and individual developers to access and build upon cutting-edge AI technology.

Conclusion

By combining strong performance, cost-effectiveness, and open-source accessibility, DeepSeek R1 represents a significant development in the AI landscape, challenging the dominance of proprietary models and potentially accelerating innovation in the field.

References:

- DeepSeek-R1 – GitHub: https://github.com/deepseek-ai/DeepSeek-R1/forks
- Money control - https://www.moneycontrol.com/news/opinion/deepseek-r1-and-the-open-source-ai-revolution-what-it-means-for-india-12920902.html
- GeeksforGeeks - https://www.geeksforgeeks.org/deepseek-r1-technical-overview-of-its-architecture-and-innovations/
- https://www.youtube.com/watch?v=KTonvXhsxpc
- Aws. Amazon - https://aws.amazon.com/what-is/reinforcement-learning/
- Cloud. Google - https://cloud.google.com/blog/products/ai-machine-learning/supervised-fine-tuning-for-gemini-llm
- Statista - https://www.statista.com/chart/33839/prices-for-processing-one-million-input-output-tokens-on-different-ai-models/