

Grok-1

Grok-1 is a mixture-of-experts (MoE) large language model (LLM) with 314B parameters which includes the open release of the base model weights and network architecture.

Grok-1 is trained by xAI and consists of MoE model that activates 25% of the weights for a given token at inference time. The pretraining cutoff date for Grok-1 is October 2023.

As stated in the [official announcement](#), Grok-1 is the raw base model checkpoint from the pre-training phase which means that it has not been fine-tuned for any specific application like conversational agents.

The model has been [released](#) under the Apache 2.0 license.

Results and Capabilities

According to the initial [announcement](#), Grok-1 demonstrated strong capabilities across reasoning and coding tasks. The last publicly available results show that Grok-1 achieves 63.2% on the HumanEval coding task and 73% on MMLU. It generally outperforms ChatGPT-3.5 and Inflection-1 but still falls behind improved models like GPT-4.

Benchmark	Grok-0 (33B)	LLaMa 2 70B	Inflection-1	GPT-3.5	Grok-1	Palm 2	Claude 2	GPT-4
GSM8k	56.8% 8-shot	56.8% 8-shot	62.9% 8-shot	57.1% 8-shot	62.9% 8-shot	80.7% 8-shot	88.0% 8-shot	92.0% 8-shot
MMLU	65.7% 5-shot	68.9% 5-shot	72.7% 5-shot	70.0% 5-shot	73.0% 5-shot	78.0% 5-shot	75.0% 5-shot + CoT	86.4% 5-shot
HumanEval	39.7% 0-shot	29.9% 0-shot	35.4% 0-shot	48.1% 0-shot	63.2% 0-shot	-	70% 0-shot	67% 0-shot
MATH	15.7% 4-shot	13.5% 4-shot	16.0% 4-shot	23.5% 4-shot	23.9% 4-shot	34.6% 4-shot	-	42.5% 4-shot

Grok-1 was also reported to score a C (59%) compared to a B (68%) from GPT-4 on the Hungarian national high school finals in mathematics.

Human-graded evaluation	Grok-0	GPT-3.5	Claude 2	Grok-1	GPT-4
Hungarian National High School Math Exam (May 2023)	37% 1-shot	41% 1-shot	55% 1-shot	59% 1-shot	68% 1-shot

Check out the model here: <https://github.com/xai-org/grok-1>

Due to the size of Grok-1 (314B parameters), xAI recommends a multi-GPU machine to test the model.

References

- [Open Release of Grok-1](#)
- [Announcing Grok](#)

Last updated on September 19, 2024