

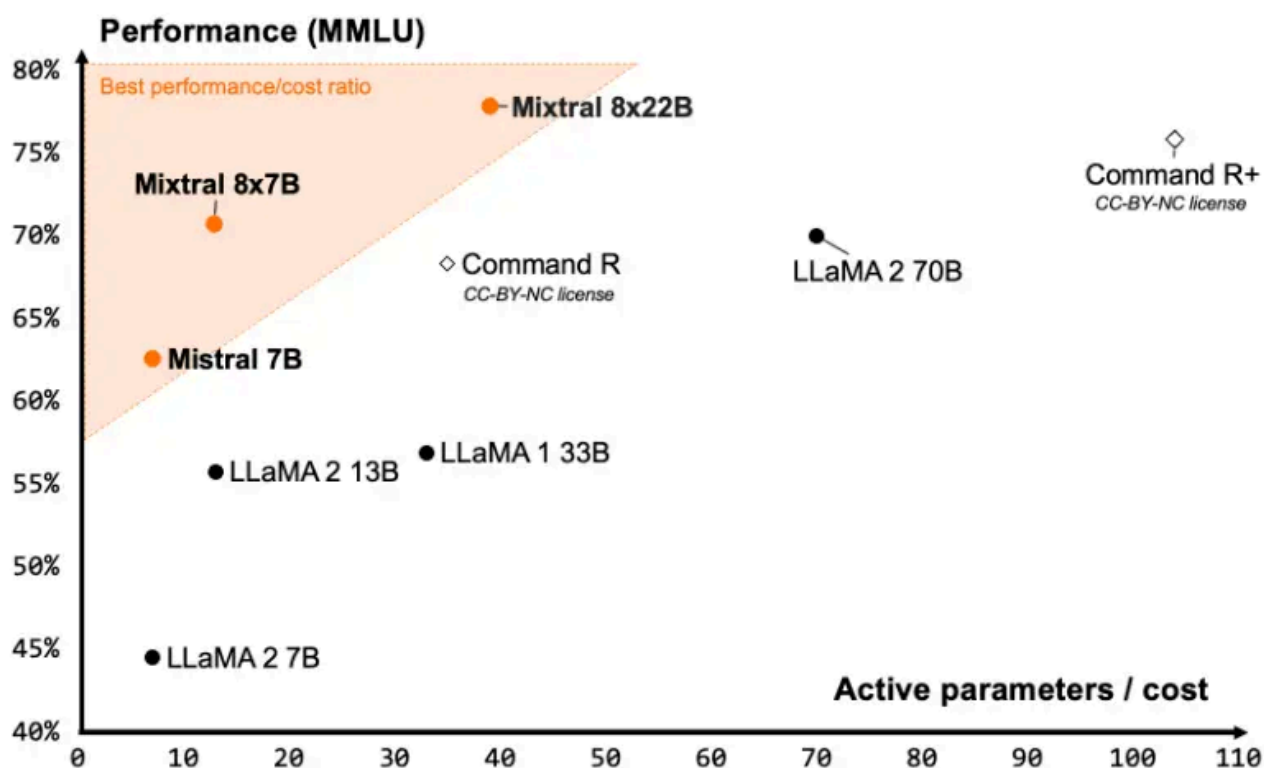
Mixtral 8x22B

Mixtral 8x22B is a new open large language model (LLM) released by Mistral AI. Mixtral 8x22B is characterized as a sparse mixture-of-experts model with 39B active parameters out of a total of 141B parameters.

Capabilities

Mixtral 8x22B is trained to be a cost-efficient model with capabilities that include multilingual understanding, math reasoning, code generation, native function calling support, and constrained output support. The model supports a context window size of 64K tokens which enables high-performing information recall on large documents.

Mistral AI claims that Mixtral 8x22B delivers one of the best performance-to-cost ratio community models and it is significantly fast due to its sparse activations.



Source: [Mistral AI Blog](#)

Results

According to the [official reported results](#), Mixtral 8x22B (with 39B active parameters) outperforms state-of-the-art open models like Command R+ and Llama 2 70B on several reasoning and knowledge benchmarks like MMLU, HellaS, TriQA, NaturalQA, among others.

Model	Active parameters	Common sense and reasoning					Knowledge	
		MMLU	HellaS	WinoG	Arc C (5)	Arc C (25)	TriQA	NaturalQS
LLaMA 2 70B	70B	69.9%	87.1%	83.2%	86.0%	85.1%	77.57%	35.5%
CC-BY-NC license	Command R	35B	68.2%	87.0%	81.5%	-	66.5%	-
	Command R+	104B	75.7%	88.6%	85.4%	-	71.0%	-
Mistral 7B	7B	62.47%	83.1%	78.0%	77.2%	78.1%	68.8%	28.1%
Mixtral 8x7B	12.9B	70.63%	86.6%	81.2%	85.8%	85.9%	78.4%	36.5%
Mixtral 8x22B	39B	77.75%	88.5%	84.7%	91.3%	91.3%	82.2%	40.1%

Source: [Mistral AI Blog](#)

Mixtral 8x22B outperforms all open models on coding and math tasks when evaluated on benchmarks such as GSM8K, HumanEval, and Math. It's reported that Mixtral 8x22B Instruct achieves a score of 90% on GSM8K (maj@8).

Model	Active parameters	Coding		Maths		
		HumanE	MBPP	GSM8K maj@1 (5-shot)	GSM8K maj@8 (8-shot)	Math maj@4
LLaMA 2 70B	70B	29.3%	49.8%	53.6%	69.6%	13.8%
CC-BY-NC license	Command R	-	-	56.6%	-	-
	Command R+	-	-	70.7%	-	-
Mistral 7B	7B	26.22%	50.2%	36.5%	50.0%	12.7%
Mixtral 8x7B	12.9B	40.2%	60.7%	58.4%	74.4%	28.4%
Mixtral 8x22B	39B	45.1%	71.2%	78.6%	88.4%	41.8%

Source: [Mistral AI Blog](#)

More information on Mixtral 8x22B and how to use it here:

https://docs.mistral.ai/getting-started/open_weight_models/#operation/listModels

The model is released under an Apache 2.0 license.

