WIKIPEDIA
The Free Encyclopedia

Mistral AI

Mistral AI is a French company specializing in artificial intelligence (AI) products. Founded in April 2023 by former employees of Meta Platforms and Google DeepMind,[1] the company has quickly risen to prominence in the AI sector.

The company focuses on producing open source large language models,[2] emphasizing the foundational importance of free and open-source software, and positioning itself as an alternative to proprietary models.[3]

In October 2023, Mistral AI raised €385 million.[4] By December 2023, it was valued at over $2 billion.[5][6][7]

In June 2024, Mistral AI announced a new funding round of €600 million ($645 million), significantly boosting its valuation to €5.8 billion ($6.2 billion).[8] This round was led by the venture capital firm General Catalyst, with participation from existing investors.[9]

Mistral AI has published three open-source models available as weights.[10] Additionally, three more models—Small, Medium, and Large—are available via API only.[11][12]

Based on valuation, the company is in fourth place in the global AI race and in first place outside the San Francisco Bay Area.[13] Mistral AI aims to "democratize" AI by focusing on open-source innovation.[14]

History

Mistral AI was co-founded in April 2023 by Arthur Mensch, Guillaume Lample and Timothée Lacroix.

Prior to co-founding Mistral AI, Arthur Mensch worked at Google DeepMind which is Google's artificial intelligence laboratory, while Guillaume Lample and Timothée Lacroix worked at Meta Platforms.[15] The co-founders met while students at École polytechnique. Mistral is named for a strong wind that blows in France.[16]

In June 2023, the start-up carried out a first fundraising of €105 million ($117 million) with investors including the American fund Lightspeed Venture Partners, Eric Schmidt, Xavier Niel and JCDecaux. The valuation is then estimated by the Financial Times at €240 million ($267 million).

On 27 September 2023, the company made its language processing model "Mistral 7B" available under the free Apache 2.0 license. This model has 7 billion parameters, a small size compared to its competitors.
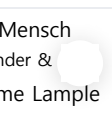
On 10 December 2023, Mistral AI announced that it had raised €385 million ($428 million) as part of its second fundraising. This round of financing notably involves the Californian fund Andreessen Horowitz, BNP Paribas and the software publisher Salesforce.[17]

On 11 December 2023, the company released the Mixtral 8x7B model with 46.7 billion parameters but using only 12.9 billion per token thanks to the mixture of experts architecture. The model masters 5 languages (French, Spanish, Italian, English and German) and outperforms, according to its developers' tests, the "LLama 2 70B" model from Meta. A version trained to follow instructions and called "Mixtral 8x7B Instruct" is also offered.[18]

On 26 February 2024, Microsoft announced a new partnership with the company to expand its presence in the rapidly evolving artificial intelligence industry. Under the agreement, Mistral's rich language models will be available on Microsoft's Azure cloud, while the multilingual conversational assistant "Le Chat" will be launched in the style of ChatGPT.[19]

On 10 April 2024, the company released the mixture of expert models, Mixtral 8x22B, offering high performance on various benchmarks compared to other open models.

Mistral AI	
	MISTRAL AI_
Company type	Private
Industry	Artificial intelligence
Founded	28 April 2023
Founders	Arthur Mensch (Co-Founder & CEO)
	Guillaume Lample (Co-Founder & Chief Scientist)
	Timothée Lacroix (Co-Founder & CTO)
Headquarters	Paris, France
Products	Mistral 7B
	Mixtral 8x7B
	Mistral Medium
	Mistral Large
	Mixtral 8x22B
	Codestral 22B
	NeMo 12B
Website	mistral.ai (https://mistral.ai/)

On 16 April 2024, reporting revealed that Mistral was in talks to raise €500 million, a deal that would more than double its current valuation to at least €5 billion.[20]

Models

Open Weight Models

Mistral 7B

Mistral 7B is a 7.3B parameter language model using the transformers architecture. Officially released on September 27, 2023, via a BitTorrent magnet link,[21] and Hugging Face.[22] The model was released under the Apache 2.0 license. The release blog post claimed the model outperforms LLaMA 2 13B on all benchmarks tested, and is on par with LLaMA 34B on many benchmarks tested.[23]

Mistral 7B uses grouped-query attention (GQA), which is a variant of the standard attention mechanism. Instead of computing attention over all the hidden states, it computes attention over groups of hidden states.[24]

Both a base model and "instruct" model were released with the latter receiving additional tuning to follow chat-style prompts. The fine-tuned model is only intended for demonstration purposes, and does not have guardrails or moderation built-in.[23]

Mixtral 8x7B

Much like Mistral's first model, Mixtral 8x7B was released via a BitTorrent link posted on Twitter on December 9, 2023,[2] and later Hugging Face and a blog post were released two days later.[18]

Unlike the previous Mistral model, Mixtral 8x7B uses a sparse mixture of experts architecture. The model has 8 distinct groups of "experts", giving the model a total of 46.7B usable parameters.[25][26] Each single token can only use 12.9B parameters, therefore giving the speed and cost that a 12.9B parameter model would incur.[18]

Mistral AI's testing shows the model beats both LLaMA 70B, and GPT-3.5 in most benchmarks.[27]

In March 2024, research conducted by Patronus AI comparing performance of LLMs on a 100-question test with prompts to generate text from books protected under U.S. copyright law found that Open AI's GPT-4, Mixtral, Meta AI's LLaMA-2, and Anthropic's Claude2 generated copyrighted text verbatim in 44%, 22%, 10%, and 8% of responses respectively.[28][29]

Mixtral 8x22B

Similar to Mistral's previous open models, Mixtral 8x22B was released via a BitTorrent link on Twitter on April 10, 2024,[30] with a release on Hugging Face soon after.[31]

Codestral 22B

Codestral is Mistral's first code focused open weight model. Codestral was launched on 29 May 2024. It is a lightweight model specifically built for code generation tasks. As of its release date, this model surpasses Meta's Llama3 70B and DeepSeek Coder 33B (78.2% - 91.6%), another code-focused model on the HumanEval FIM benchmark.[32] Mistral claims Codestral is fluent in more than 80 Programming languages[33] Codestral has its own license which forbids the usage of Codestral for Commercial purposes.[34]

Mathstral 7B

Another 7B model that Mistral released on 16 July 2024. it's focused on STEM subjects, it achieved 56.6% on MATH benchmark and 63.47% on MMLU. it was released under Apache 2.0 license.[35] it has a context length of 32k tokens. it was produced in collaboration with Project Numina.[36]

Codestral Mamba 7B

Codestral Mamba is based on Mamba 2 architecture, which allows it to generate responses even with longer input.[36] and unlike Codestral, it was released under Apache 2.0 license. only instruct version was released.[37]

API-Only Models

Unlike Mistral 7B, Mixtral 8x7B and Mixtral 8x22B, the following models are closed-source and only available through the Mistral API.[38]

Mistral Large

Mistral Large was launched on February 26, 2024, and Mistral claims it is second in the world only to OpenAI's GPT-4.

It is fluent in English, French, Spanish, German, and Italian, with Mistral claiming understanding of both grammar and cultural context, and provides coding capabilities. As of early 2024, it is Mistral's flagship AI.[39] It is also available on Microsoft Azure.

Mistral Large 2 was released in July 2024, with 123B parameters and 128k context window. It is available for free with a Mistral Research Licence, and with a commercial licence for commercial purposes.[40]

Mistral Medium

Mistral Medium is trained in various languages including English, French, Italian, German, Spanish and code with a score of 8.6 on MT-Bench.[41] It is ranked in performance above Claude and below GPT-4 on the LMSys ELO Arena benchmark.[42]

The number of parameters, and architecture of Mistral Medium is not known as Mistral has not published public information about it.

Mistral Small

Like the Large model, Small was launched on February 26, 2024. It is intended to be a light-weight model for low latency, with better performance than Mixtral 8x7B.[43]

References

1. "France's unicorn start-up Mistral AI embodies its artificial intelligence hopes" (https://www.lemonde.fr/en/economy/article/2023/12/12/french-unicorn-start-up-mistral-ai-embodies-its-artificial-intelligence-hopes_6337125_19.html). Le Monde.fr. 2023-12-12. Retrieved 2023-12-16.

2. "Buzzy Startup Just Dumps AI Model That Beats GPT-3.5 Into a Torrent Link" (https://gizmodo.com/mistral-artificial-intelligence-gpt-3-openai-1851091217). Gizmodo. 2023-12-12. Retrieved 2023-12-16.

3. "Bringing open AI models to the frontier" (https://mistral.ai/news/about-mistral-ai/). Mistral AI. 27 September 2023. Retrieved 4 January 2024.

4. Metz, Cade (10 December 2023). "Mistral, French A.I. Start-Up, Is Valued at $2 Billion in Funding Round" (https://www.nytimes.com/2023/12/10/technology/mistral-ai-funding.html). The New York Times.

5. Fink, Charlie. "This Week In XR: Epic Triumphs Over Google, Mistral AI Raises $415 Million, $56.5 Million For Essential AI" (https://www.forbes.com/sites/charliefink/2023/12/14/this-week-in-xr-epic-triumphs-over-google-mistral-ai-raises-415-million-565-million-for-essential-ai/). Forbes. Retrieved 2023-12-16.

6. "A French AI start-up may have commenced an AI revolution, silently" (https://www.hindustantimes.com/business/a-french-ai-start-up-may-have-commenced-an-ai-revolution-silently-101702370816617.html). Hindustan Times. December 12, 2023.

7. "French AI start-up Mistral secures €2bn valuation" (https://www.ft.com/content/ea29ddf8-91cb-45e8-86a0-f501ab7ad9bb). ft.com Financial Times.

8. Kharpal, Arjun (2024-05-24). "CEOs of AI startups backed by Microsoft and Amazon are the new tech rockstars" (https://www.cnbc.com/2024/05/24/mistral-ai-anthropic-ceos-are-the-new-tech-rockstars-in-ai-boom.html). *CNBC*. Retrieved 2024-06-13.

9. "Tripling Down on Mistral AI | General Catalyst" (https://www.generalcatalyst.com/perspectives/tripling-down-on-mistral-ai). *www.generalcatalyst.com*. Retrieved 2024-06-13.

10. "Open-weight models and Mistral AI Large Language Models" (https://docs.mistral.ai/models/). docs.mistral.ai. Retrieved 2024-01-04.

11. "Endpoints and Mistral AI Large Language Models" (https://docs.mistral.ai/platform/endpoints/#medium). docs.mistral.ai.

12. "Endpoints and benchmarks | Mistral AI Large Language Models" (https://docs.mistral.ai/platform/endpoints/). *docs.mistral.ai*. Retrieved 2024-03-06.

13. Bratton, Laura (2024-06-12). "OpenAI's French rival Mistral AI is now worth $6 billion. That's still a fraction of its top competitors" (https://qz.com/openai-mistral-ai-funding-valuation-microsoft-1851535049). Quartz (publication). Retrieved 2024-06-13.

14. Webb, Maria (2024-01-02). "Mistral AI: Exploring Europe's Latest Tech Unicorn" (https://www.techopedia.com/mistral-ai-exploring-europes-latest-tech-unicorn). *techopedia.com*. Retrieved 2024-06-13.

15. "France's unicorn start-up Mistral AI embodies its artificial intelligence hopes" (https://www.lemonde.fr/en/economy/article/2023/12/12/french-unicorn-start-up-mistral-ai-embodies-its-artificial-intelligence-hopes_6337125_19.html). *Le Monde.fr*. 12 December 2023.

16. Journal, Sam Schechner | Photographs by Edouard Jacquinet for The Wall Street. "The 9-Month-Old AI Startup Challenging Silicon Valley's Giants" (https://www.wsj.com/tech/ai/the-9-month-old-ai-startup-challenging-silicon-valleys-giants-ee2e4c48). *WSJ*. Retrieved 2024-03-31.

17. "Mistral lève 385 M€ et devient une licorne française - le Monde Informatique" (https://www.lemondeinformatique.fr/actualites/lire-mistral-leve-385-meteuro-et-devient-une-licorne-francaise-92392.html). 11 December 2023.

18. "Mixtral of experts" (https://mistral.ai/news/mixtral-of-experts/). *mistral.ai*. 2023-12-11. Retrieved 2024-01-04.

19. Bableshwar (2024-02-26). "Mistral Large, Mistral AI's flagship LLM, debuts on Azure AI Models-as-a-Service" (https://techcommunity.microsoft.com/t5/ai-machine-learning-blog/mistral-large-mistral-ai-s-flagship-llm-debuts-on-azure-ai/ba-p/4066996). techcommunity.microsoft.com. Retrieved 2024-02-26.

20. "Mistral in talks to raise €500mn at €5bn valuation" (https://www.ft.com/content/358fc1d8-8276-4e81-9ee6-e0391575d569). *www.ft.com*. Retrieved 2024-04-19.

21. Goldman, Sharon (2023-12-08). "Mistral AI bucks release trend by dropping torrent link to new open source LLM" (https://venturebeat.com/ai/mistral-ai-bucks-release-trend-by-dropping-torrent-link-to-new-open-source-llm/). *VentureBeat*. Retrieved 2024-01-04.

22. Coldewey, Devin (27 September 2023). "Mistral AI makes its first large language model free for everyone" (https://techcrunch.com/2023/09/27/mistral-ai-makes-its-first-large-language-model-free-for-everyone/). *TechCrunch*. Retrieved 4 January 2024.

23. "Mistral 7B" (https://mistral.ai/news/announcing-mistral-7b/). *mistral.ai*. Mistral AI. 27 September 2023. Retrieved 4 January 2024.

24. Jiang, Albert Q.; Sablayrolles, Alexandre; Mensch, Arthur; Bamford, Chris; Chaplot, Devendra Singh; Casas, Diego de las; Bressand, Florian; Lengyel, Gianna; Lample, Guillaume (2023-10-10). "Mistral 7B". arXiv:2310.06825v1 (https://arxiv.org/abs/2310.06825v1) [cs.CL (https://arxiv.org/archive/cs.CL)].

25. "Mixture of Experts Explained" (https://huggingface.co/blog/moe). *huggingface.co*. Retrieved 2024-01-04.

26. Marie, Benjamin (2023-12-15). "Mixtral-8x7B: Understanding and Running the Sparse Mixture of Experts" (https://towardsdatascience.com/mixtral-8x7b-understanding-and-running-the-sparse-mixture-of-experts-0e3fc7fde818). *Medium*. Retrieved 2024-01-04.

27. Franzen, Carl (2023-12-11). "Mistral shocks AI community as latest open source model eclipses GPT-3.5 performance" (https://venturebeat.com/ai/mistral-shocks-ai-community-as-latest-open-source-model-eclipses-gpt-3-5-performance/). *VentureBeat*. Retrieved 2024-01-04.

28. Field, Hayden (March 6, 2024). "Researchers tested leading AI models for copyright infringement using popular books, and GPT-4 performed worst" (https://www.cnbc.com/2024/03/06/gpt-4-researchers-tested-leading-ai-models-for-copyright-infringement.html). CNBC. Retrieved March 6, 2024.

29. "Introducing CopyrightCatcher, the first Copyright Detection API for LLMs" (https://www.patronus.ai/blog/introducing-copyright-catcher). Patronus AI. March 6, 2024. Retrieved March 6, 2024.

30. @MistralAI (April 10, 2024). "Torrent" (https://x.com/MistralAI/status/1777869263778291896) (Tweet) – via Twitter.

31. "mistralai/Mixtral-8x22B-v0.1 · Hugging Face" (https://huggingface.co/mistralai/Mixtral-8x22B-v0.1). *huggingface.co*. Retrieved 2024-05-05.

32. AI, Mistral (2024-05-29). "Codestral: Hello, World!" (https://mistral.ai/news/codestral/). *mistral.ai*. Retrieved 2024-05-30.

33. Sharma, Shubham (2024-05-29). "Mistral announces Codestral, its first programming focused AI model" (https://venturebeat.com/ai/mistral-announces-codestral-its-first-programming-focused-ai-model/). *VentureBeat*. Retrieved 2024-05-30.

34. Wiggers, Kyle (2024-05-29). "Mistral releases Codestral, its first generative AI model for code" (https://techcrunch.com/2024/05/29/mistral-releases-its-first-generative-ai-model-for-code/). *TechCrunch*. Retrieved 2024-05-30.

35. AI, Mistral (2024-07-16). "MathΣtral" (https://mistral.ai/news/mathstral/). *mistral.ai*. Retrieved 2024-07-16.

36. David, Emilia (2024-07-16). "Mistral releases Codestral Mamba for faster, longer code generation" (https://venturebeat.com/ai/mistral-releases-codestral-mamba-for-faster-longer-code-generation/). *VentureBeat*. Retrieved 2024-07-17.

37. AI, Mistral (2024-07-16). "Codestral Mamba" (https://mistral.ai/news/codestral-mamba/). *mistral.ai*. Retrieved 2024-07-16.

38. "Pricing and rate limits | Mistral AI Large Language Models" (https://docs.mistral.ai/platform/pricing/). *docs.mistral.ai*. Retrieved 2024-01-22.

39. AI, Mistral (2024-02-26). "Au Large" (https://mistral.ai/news/mistral-large/). *mistral.ai*. Retrieved 2024-03-06.

40. AI, Mistral (2024-07-24). "Large Enough" (https://mistral.ai/news/mistral-large-2407/). *mistral.ai*. Retrieved 2024-07-24.

41. AI, Mistral (2023-12-11). "La plateforme" (https://mistral.ai/news/la-plateforme/). *mistral.ai*. Retrieved 2024-01-22.

42. "LMSys Chatbot Arena Leaderboard - a Hugging Face Space by lmsys" (https://huggingface.co/spaces/lmsys/chatbot-arena-leaderboard). *huggingface.co*. Retrieved 2024-01-22.

43. AI, Mistral (2024-02-26). "Au Large" (https://mistral.ai/news/mistral-large/). *mistral.ai*. Retrieved 2024-03-06.

External links

- Official website (https://mistral.ai/)