

Google releases Gemini 2.5 Pro, a thinking model that debuts first on LMArena

Source copy kept by wheresthe.ai for the story "Google releases Gemini 2.5 Pro, a thinking model that debuts first on LMArena" (<https://wheresthe.ai/d/google-gemini-2-5-pro-2025-03/>).

Source: <https://blog.google/innovation-and-ai/models-and-research/google-deepmind/gemini-model-thinking-updates-march-2025/>

Text saved 18 Oct 2026 (the page as published).

SHA-256 of the saved text: f47c2e709f9e931114d7cc7e7efa0fcf96858212bbec254cbfeb7bee46317e32

Gemini 2.5: Our most intelligent AI model

Mar 25, 2025

|

x.com

Facebook

LinkedIn

Mail

Copy link

Gemini 2.5 is a thinking model, designed to tackle increasingly complex problems. Our first 2.5 model, Gemini 2.5 Pro Experimental, leads common benchmarks by meaningful margins and showcases strong reasoning and code capabilities.

Koray Kavukcuoglu

CTO of Google DeepMind

Share

x.com

Facebook

LinkedIn

Mail

Copy link

In this article

Last updated March 26 Today we're introducing Gemini 2.5, our most intelligent AI model. Our first 2.5 release is an experimental version of 2.5 Pro, which is state-of-the-art on a wide range of benchmarks and debuts at #1 on LMArena by a significant margin. Gemini 2.5 models are thinking models, capable of reasoning through their thoughts before responding, resulting in enhanced performance and improved accuracy. In the field of AI, a system's capacity for "reasoning" refers to more than just classification and prediction. It refers to its ability to analyze information, draw logical conclusions, incorporate context and nuance, and make informed decisions. For a long time, we've explored ways of making AI smarter and more capable of reasoning through techniques like reinforcement learning and chain-of-thought prompting . Building on this, we recently introduced our first thinking model, Gemini 2.0 Flash Thinking . Now, with Gemini 2.5, we've achieved a new level of performance by combining a significantly enhanced base model with improved post-training. Going forward, we're building these thinking capabilities directly into all of our models, so they can handle more complex problems and support even more capable, context-aware agents.

Introducing Gemini 2.5 Pro Gemini 2.5 Pro Experimental is our most advanced model for complex tasks. It tops the LMArena leaderboard — which measures human preferences — by a significant margin, indicating a highly

capable model equipped with high-quality style. 2.5 Pro also shows strong reasoning and code capabilities, leading on common coding, math and science benchmarks. Gemini 2.5 Pro is available now in Google AI Studio and in the Gemini app for Gemini Advanced users, and will be coming to Vertex AI soon. We'll also introduce pricing in the coming weeks, enabling people to use 2.5 Pro with higher rate limits for scaled production use.

Updated March 26 with new MRCR (Multi Round Coreference Resolution) evaluations

Enhanced reasoning Gemini 2.5 Pro is state-of-the-art across a range of benchmarks requiring advanced reasoning. Without test-time techniques that increase cost, like majority voting, 2.5 Pro leads in math and science benchmarks like GPQA and AIME 2025. It also scores a state-of-the-art 18.8% across models without tool use on Humanity's Last Exam, a dataset designed by hundreds of subject matter experts to capture the human frontier of knowledge and reasoning.

Advanced coding We've been focused on coding performance, and with Gemini 2.5 we've achieved a big leap over 2.0 — with more improvements to come. 2.5 Pro excels at creating visually compelling web apps and agentic code applications, along with code transformation and editing. On SWE-Bench Verified, the industry standard for agentic code evals, Gemini 2.5 Pro scores 63.8% with a custom agent setup. Here's an example of how 2.5 Pro can use its reasoning capabilities to create a video game by producing the executable code from a single line prompt.

Building on the best of Gemini Gemini 2.5 builds on what makes Gemini models great — native multimodality and a long context window. 2.5 Pro ships today with a 1 million token context window (2 million coming soon), with strong performance that improves over previous generations. It can comprehend vast datasets and handle complex problems from different information sources, including text, audio, images, video and even entire code repositories. Developers and enterprises can start experimenting with Gemini 2.5 Pro in Google AI Studio now, and Gemini Advanced users can select it in the model dropdown on desktop and mobile. It will be available on Vertex AI in the coming weeks. As always, we welcome feedback so we can continue to improve Gemini's impressive new abilities at a rapid pace, all with the goal of making our AI more helpful.

Posted in:

Related stories

Google DeepMind

TIME has named Google SynthID one of the Best Inventions of 2026.

Google DeepMind

We're powering European scientific discovery with advanced AI compute and frontier models.

By

Maureen Costello

Safety & Security

Developing Gemini 4 Argon responsibly and safely

By

Anca Dragan

& Raluca Ada Popa

Developer tools

Gemini 4 Argon is now available for developers.

Gemini models

Gemini 4 Argon: Built to solve your hardest problems

By

Koray Kavukcuoglu

Health

Applying technology to understand and treat disease

By

Yossi Matias

& Pushmeet Kohli