# Google’s Gemini 2.5 leads AI race with record-breaking token capacity and benchmark scores



Google’s Gemini 2.5 has made waves in artificial intelligence, recently topping various benchmarks that gauge its capabilities across multiple domains. This latest offering from the tech giant has outperformed competitors on coding tests, earning a leading position on the WebDev Arena leaderboard, which ranks AI models predominantly used for coding tasks. Such performances indicate a tightening race among AI models, especially against formidable foes like Claude and ChatGPT 4.

A closer look at Gemini 2.5 unveils its remarkable capabilities. Beyond securing the top spot in coding functionalities, it also excelled in creative writing and style control. Particularly noteworthy is its performance in standardized IQ tests, where it notably achieved an IQ score of 124 on the Mensa Norway test. In contrast, its performance dipped slightly in offline mode, recording a score of 115, thereby placing it in a tie for second with OpenAI’s ChatGPT.

The model further demonstrates its prowess scoring 86.7% on the AIME 2025 math test and 84% on the GPQA science assessment. Interestingly, despite a low score of 18.8% on Humanity’s Last Exam, Gemini 2.5 still emerged victorious, surpassing both Claude and OpenAI’s latest model, o3. This success can be attributed to its extensive context window, capable of processing up to 1 million tokens, far exceeding the 128K tokens that Claude and ChatGPT can handle. Google's vision reportedly includes plans to expand this window to an impressive 2 million tokens, indicating an ongoing commitment to enhancing user experience and performance.

In a statement following its commercial release, Google noted, “Gemini 2.5 models are thinking models, capable of reasoning through their thoughts before responding, resulting in enhanced performance and improved accuracy.” Such advancements are not only theoretical – a pro version offers additional features, with pricing starting as low as $2.50 for input and $15.00 for output, making it an economically attractive choice compared to its competitors.

The implications of these advancements stretch far beyond test scores, however. AI chatbots, including those powered by Gemini 2.5, are transforming workplace efficiency by introducing heightened levels of automation and personalisation. Governments, too, are exploring AI’s potential to enhance public services, although this trend does not come without concerns regarding potential misuse, notably in areas like censorship and disinformation.

Moreover, the rise of blockchain technology is set to intersect significantly with AI development. Recent innovations, such as the launch of Space and Time’s mainnet, highlight a burgeoning collaboration that leverages zero-knowledge proofs to enhance data transparency and security across digital asset service providers. This development underlines a varied application spectrum, from decentralised finance (DeFi) to gaming, showcasing how AI and blockchain can collectively inform more secure and efficient systems.

As these technologies evolve, researchers are also actively examining the interplay between AI and blockchain, particularly as a means to combat misinformation and biases that can proliferate in AI-generated content. By integrating enterprise blockchain systems, there lies the potential for ensuring data integrity and ownership, critical factors needed for AI to function ethically and effectively in an increasingly complex landscape.

The continuous development and competition among models such as Gemini 2.5 and other leading AIs pose promising prospects. The efficacy of these systems brings fresh opportunities for innovation across various sectors, with significant implications for how we engage with technology moving forward.

As we stand at the forefront of these advancements, the intersection of AI, blockchain, and real-world application is poised to redefine the technological narrative.

### Reference Map

1. Paragraphs 1-4, 6
2. Paragraphs 2, 4
3. Paragraphs 2, 3
4. Paragraph 5
5. Paragraph 5
6. Paragraphs 5, 6
7. Paragraph 6

Source: [Noah Wire Services](https://www.noahwire.com)

## Bibliography

1. <https://coingeek.com/google-gemini-2-5-ranks-first-in-coding-charts-ai-iq-tests/> - Please view link - unable to able to access data
2. <https://deepmind.google/technologies/gemini/> - DeepMind's official page on Gemini 2.5 Pro highlights its state-of-the-art performance across various benchmarks, including reasoning, coding, and long-context tasks. The model demonstrates significant improvements over previous versions and competitors, showcasing its advanced capabilities in AI applications.
3. <https://www.datacamp.com/blog/gemini-2-5-pro> - DataCamp's article provides an in-depth analysis of Gemini 2.5 Pro's performance across multiple benchmarks. It compares the model's results with competitors like Claude 3.7 Sonnet and OpenAI's o3-mini, highlighting Gemini 2.5 Pro's strengths in reasoning, coding, and long-context tasks.
4. <https://www.tomsguide.com/computing/gemini-2-5-pro-is-now-free-to-all-users-in-surprise-move> - Tom's Guide reports on Google's unexpected decision to make Gemini 2.5 Pro free to all users. Initially available only to paid subscribers, the model is now accessible to a broader audience, offering advanced AI capabilities without a subscription fee.
5. <https://www.analyticsvidhya.com/blog/2025/03/gemini-2-5-pro-vs-gpt-4-5/> - Analytics Vidhya compares Gemini 2.5 Pro with OpenAI's GPT-4.5 across various benchmarks. The article highlights Gemini 2.5 Pro's superior performance in reasoning, science, mathematics, and coding tasks, positioning it as a strong competitor in the AI field.
6. <https://huggingface.co/blog/lynn-mikami/gemini-2-5-pro-preview> - Hugging Face's blog post previews Gemini 2.5 Pro's performance, emphasizing its leading position in AI coding benchmarks. The article discusses the model's strengths in code generation, code editing, and agentic coding tasks, showcasing its advanced capabilities.
7. <https://medium.com/data-science-in-your-pocket/google-gemini-2-5-pro-the-best-llm-ever-172d0665336b> - Mehul Gupta's Medium article delves into Gemini 2.5 Pro's performance, highlighting its top scores in various benchmarks. The piece discusses the model's strengths in reasoning, mathematics, and coding tasks, positioning it as a leading AI model.