

Executive Summary

GLM-4.5: A Breakthrough Open-Source AI Model Excelling Across Agentic, Reasoning, and Coding Tasks

The Problem

Current AI language models face a critical limitation: while some excel at specific capabilities like mathematical reasoning or coding, no single open-source model demonstrates comprehensive excellence across all three essential domains for advanced AI systems: agentic abilities (interacting with tools and environments), complex reasoning (solving multi-step problems), and advanced coding (tackling real-world software engineering). This forces developers to choose between specialized models or settle for compromised performance.

The Breakthrough

GLM-4.5 introduces a revolutionary **hybrid reasoning approach** that combines both thinking and direct response modes in a single Mixture-of-Experts (MoE) architecture. By employing expert model iteration and multi-stage reinforcement learning during post-training, GLM-4.5 achieves remarkable performance parity with leading proprietary models while using significantly fewer parameters—only half the parameters of DeepSeek-R1 and one-third of Kimi K2.

How It Works

The model leverages a **355B parameter MoE architecture** with only 32B activated parameters for efficiency, combined with innovative training techniques including difficulty-based curriculum learning and dynamic sampling temperature adjustment. This enables GLM-4.5 to achieve **91.0% accuracy on AIME 24** (one of the most challenging mathematics competitions) and **70.1% on TAU-Bench** (agentic task evaluation), while maintaining strong coding performance at **64.2% on SWE-bench Verified**.

Why This Matters

This breakthrough democratizes access to state-of-the-art AI capabilities, enabling researchers, startups, and developers to build sophisticated applications that previously required expensive proprietary APIs. The model's open-source nature and strong performance across all three domains make it ideal for developing AI agents, automated coding assistants, and advanced reasoning systems.

The Business Opportunity

GLM-4.5 creates new possibilities for cost-effective AI products that can handle complex multi-step tasks—from autonomous software development to intelligent tutoring systems—while reducing infrastructure costs by 2-3x compared to larger proprietary models through its efficient MoE architecture.