

# VibeThinker-3B: Frontier Math and Coding Reasoning in a 3B Model

Kabui, Charles

2026-06-29

---

[Read at ToKnow.ai](#)

---

**VibeThinker-3B: Frontier Reasoning in a 3B Model**

A pocket model that matches giants on math and code

<b>3B</b> Matches Gemini 3 Pro, far larger	<b>94.3</b> AIME26 math, 97.1 with scaling	<b>96.1%</b> Pass rate on unseen LeetCode
---	---	--

June 29, 2026

ToKnow.ai

WeiboAI, the research lab inside Chinese social platform Weibo, released [VibeThinker-3B](#), a 3-billion-parameter model built on Qwen2.5-Coder-3B and tuned for verifiable reasoning, meaning problems with a checkable answer like math and code. It scores 94.3 on [AIME26](#), a hard math competition, 80.2 on LiveCodeBench v6 (solved on the first try), and a 96.1% acceptance rate on LeetCode contests it had never seen. The [technical report](#) shows it matches

or beats DeepSeek V3.2, GLM-5, and Gemini 3 Pro, models hundreds of times larger. A test-time method that spends extra compute checking each step lifts AIME26 to 97.1. A 93.4 on IFEval confirms the reasoning push did not break instruction following.

Frontier-level math and coding now fit on a laptop instead of a server rack. The earlier 1.5B version cost just \$7,800 to post-train, against \$294K for DeepSeek R1 and \$535K for MiniMax-M1, a 30x to 60x drop. A student, a small startup, or a hospital can run a top reasoning model locally, cheaply, and privately, with weights released under the permissive MIT license.

The team argues reasoning packs into a small core, while broad world knowledge still needs scale. That points to two model sizes for two jobs, not one giant model for everything. It echoes NVIDIA's [Nemotron-Cascade 2](#), which hit math-olympiad gold with 3B active parameters.

Sources:

- [VibeThinker-3B Technical Report \(arXiv\)](#)
- [VibeThinker-3B on Hugging Face](#)
- [VibeThinker GitHub Repository](#)
- [Technical Report on Hugging Face](#)

---

*Disclaimer: For information only. Accuracy or completeness not guaranteed. Illegal use prohibited. Not professional advice or solicitation. Read more: [/terms-of-service](#)*