

Mass market AI tools failed 85% of investment tasks: Report

A study conducted by investment platform DeepVest found that popular AI tools often produced incorrect calculations or 'hallucinated' when given investment management-related tasks.

BY **ALEC RICH**

Enterprise artificial intelligence (AI) assistants aren't getting the job done when it comes to financial analysis.

That's according to a new [report](#) from AI-focused investment platform DeepVest, which has released the findings of a six-week study measuring the investment workflow capabilities of several notable AI assistants and chatbot platforms. Those included OpenAI's ChatGPT, Anthropic's Claude, Perplexity, xAI's SuperGrok and Google's Gemini.

Most notably, DeepVest found that the general purpose AI tools failed a combined 85% of investment tasks assigned to them, 'often producing incorrect calculations, hallucinated data, or no results at all.'

'The results highlight a core challenge facing financial professionals today: while frontier models can appear sophisticated and persuasive, they frequently fail to



Active ETFs

RONB BCGD

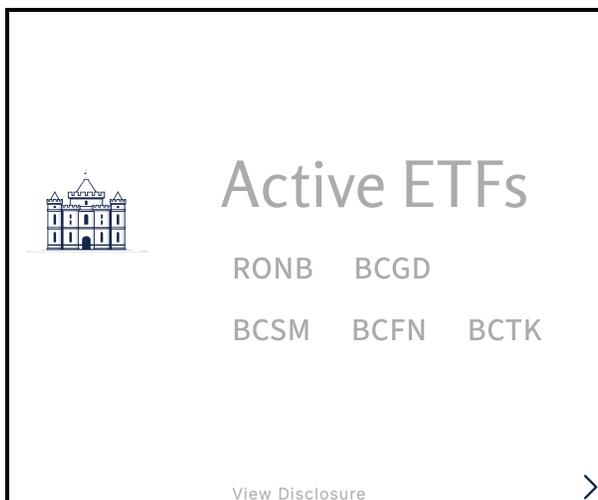
BCSM BCFN BCTK

[View Disclosure](#)



DUE DILIGENCE REPORTS

compute reliably, lack access to trustworthy financial datasets and cannot guarantee repeatable outputs,' the report states.



Active ETFs

RONB BCGD
BCSM BCFN BCTK

[View Disclosure](#) >

The tools were evaluated across ten 'investment use cases,' including rolling correlations, portfolio back testing, stock screening, macro stress testing, options analysis, portfolio optimization and maximum drawdowns.

The study noted that computational problems became more pronounced as the complexity of tasks rose. Though advisors can always check given outputs from chatbots, the report notes that such manual oversight somewhat defeats the purpose of automating tasks to begin with.

Some of the most glaring examples of issues DeepVest found included Gemini's maximum drawdowns being off by between 13% and 50%, ChatGPT attempting to pull Bitcoin prices from the Federal Reserve Economic Database, which does not track crypto and none of the platforms providing sufficient audit trails needed for regulatory compliance.



Core & Core Plus
STRENGTH IN STRUCTURE

Due Diligence Report: Core & Core Plus

Discover how portfolio design, credit quality and flexibility combine to create lasting resilience through changing market conditions.

'The results across all investment use cases show a consistent and clear pattern: frontier models consistently failed to reliably perform real investment analysis tasks that require accurate data retrieval, precise computation, and repeatable outputs,' the report states.

The report concludes that financial advisors should be using AI systems specifically designed for regulated investment workflows with 'verified data, audit trails, and deterministic outputs,' moving forward.

DeepVest is itself one of those tools competing in the portfolio analysis space. The study compared the results of the various chatbots relative to DeepVest's own platform, along with independent analysis using coding platform Python and other institutional sources like the Bloomberg Terminal.

The research was conducted in January and February using the most recently available versions of the AI assistants, DeepVest said, though it provided a disclaimer that prompt changes or newer updates could alter results.

The research comes on the heels of enterprise AI firms more formally entering the RIA space via partnerships with technology vendors and broker-dealers. For example, Anthropic on Tuesday announced the rollout of wealth management-specific plug-in tools.

Brokerage giant LPL Financial [revealed](#) that it intends to utilize the new plug-ins as part of a burgeoning partnership with Anthropic, while RIA tech giant Orion has already utilized Anthropic to support the internal architecture of its Denali AI tool.



Active ETFs

RONB BCGD

BCSM BCFN BCTK

[View Disclosure](#)



Got a news tip? Contact us anonymously at tips@citywireusa.com

Citywire AI

Your trusted research partner.

How can I avoid AI hallucinations and look-ahead bias in my investment decisions?

What caused the recent sell-off in wealth management stocks?

Why does an elite AI expert shun AI for stock selection?

What is needed for future autonomous AI agents to achieve true reasoning?

How is investor comfort with fully AI-driven advice increasing globally?

Or type your own question below.

Your prompt



GO

Related News