

Gemini 2.5 có khả năng thực hiện các thao tác tương tự con người

09:26 12/10/2025

Google mới đây ra mắt mô hình AI Gemini mới, có khả năng thực hiện các thao tác tương tự con người.



Mô hình AI Gemini 2.5 Computer Use có thể thao tác như con người

Trong thông báo mới nhất, Google giới thiệu Gemini 2.5 Computer Use – mô hình AI có khả năng thao tác trực tiếp trên các trang web và ứng dụng như con người. Mô hình này được thiết kế để hiểu yêu cầu, quan sát giao diện, sau đó tự thực hiện các tác vụ trong môi trường được tạo cho người dùng thật. Nhờ khả năng hiểu ngữ cảnh và lập luận trực quan, Gemini 2.5 có thể xử lý các nhiệm vụ phức tạp mà không cần giao diện đặc biệt cho AI.

Theo thông tin từ Google thì Gemini 2.5 Computer Use có thể được sử dụng để kiểm tra giao diện người dùng hoặc điều hướng các giao diện được thiết kế cho những người không có API hoặc kết nối trực tiếp khác.

Các phiên bản khác của mô hình này đã được sử dụng trong các tính năng tác nhân (agentic) của AI Mode và Project Mariner- một nguyên mẫu nghiên cứu

cho phép các tác nhân AI tự động thực hiện tác vụ trong trình duyệt, chẳng hạn như thêm sản phẩm vào giỏ hàng dựa trên danh sách mua sắm.

Thông báo của Google được đưa ra chỉ một ngày sau khi OpenAI công bố loạt ứng dụng ChatGPT mới tại sự kiện Dev Day thường niên. Động thái này cho thấy Google không muốn đứng ngoài cuộc đua về các tác nhân AI (ChatGPT Agent) – công nghệ cho phép trí tuệ nhân tạo tự động thực hiện các nhiệm vụ phức tạp thay cho người dùng. Trong khi đó, đối thủ Anthropic đã giới thiệu tính năng computer use trong mô hình Claude từ năm ngoái.

Trong các video demo được công bố, Google trình diễn cách Gemini 2.5 Computer Use vận hành và nhấn mạnh rằng hiệu suất đã được cải thiện gấp ba lần so với trước đây. Hãng cũng khẳng định mô hình này vượt trội hơn các đối thủ hàng đầu trên nhiều bài kiểm tra chuẩn về hiệu năng web và di động.

Tuy nhiên, khác với ChatGPT Agent của OpenAI hay computer use của Anthropic, mô hình mới của Google chỉ có quyền truy cập vào trình duyệt web thay vì toàn bộ môi trường máy tính. Hãng cũng lưu ý rằng mô hình hiện chưa được tối ưu hóa để điều khiển ở cấp độ hệ điều hành máy tính để bàn, và mới chỉ hỗ trợ 13 thao tác cơ bản như mở trình duyệt, nhập văn bản, kéo – thả các thành phần.

Hiện tại, Gemini 2.5 Computer Use đã được cung cấp cho các nhà phát triển thông qua Google AI Studio và Vertex AI, đồng thời có phiên bản demo trên Browserbase – nơi người dùng có thể quan sát trực tiếp quá trình AI hoàn thành các tác vụ.

Hà Trần (t/h)