

رتبه ششم جهان با ۸۳.۳ امتیاز؛ شیائومی با MiMo-V۲.۶-Pro صدرنشین مدل‌های وزن‌باز شد

خانواده مدل‌های هوش مصنوعی MiMo-V۲.۶ با عملکردی خیره‌کننده در یک آزمون از هوش مصنوعی GPT-۵ پیشی گرفت.

به گزارش روابط عمومی دانشگاه سئوی ایران، شیائومی خانواده‌ی MiMo-V۲.۶ را با دو مدل Pro و Flash منتشر کرد؛ مدل‌هایی که تمرکز اصلی آن‌ها روی استدلال، عامل‌های هوش مصنوعی، برنامه‌نویسی و انجام کارهای چندمرحله‌ای است. شیائومی می‌گوید MiMo-V۲.۶-Pro در بسیاری از بنچمارک‌های عامل‌محور عملکردی در سطح مدل‌هایی مانند Claude Opus ۵ و GPT-۵.۶ Sol دارد، درحالی‌که وزن‌های آن به‌صورت عمومی منتشر شده‌اند.

در یکی از رتبه‌بندی‌های منتشرشده توسط World of AI Bench، هوش مصنوعی MiMo-V۲.۶-Pro امتیاز ترکیبی ۸۳/۳ به دست آورد و در رتبه‌ی ششم قرار گرفت؛ فاصله‌اش با Claude Opus ۵ حدود ۰/۴ امتیاز اعلام شد و اندکی بالاتر از GPT-۵.۶ Sol قرار گرفت. این بنچمارک یک ارزیابی مستقل و نه استاندارد رسمی مورد استفاده‌ی همه‌ی شرکت‌های هوش مصنوعی محسوب می‌شود.

ارزیابی مستقل دیگری از Artificial Analysis نیز امتیاز ۴۶/۳۲ را برای MiMo-V۲.۶-Pro ثبت کرده و آن را در صدر مدل‌های وزن‌باز قرار داد. این مدل پنجره‌ی کانتکست یک میلیون توکنی دارد و از ورودی متن، تصویر، ویدئو و صدا پشتیبانی می‌کند.

یکی از نکات مهم MiMo-V۲.۶، تمرکز روی یادگیری تقویتی در مقیاس بالا است. شیائومی می‌گوید آموزش RL دو مدل در کمتر از ۶ روز و طی ۳۰ مرحله انجام شد. هزینه‌ی آموزش نسخه‌ی Pro حدود ۲/۶۲ میلیون دلار اعلام شده است. DeepSWE v۱.۱ نیز در طول این فرایند برای Pro از ۵۸/۴ به ۷۲/۶ رسید.

نکته‌ی مهم دیگر، قیمت است. شیائومی اعلام کرده قیمت API این نسل نسبت به MiMo-V۲.۵ افزایش پیدا نکرده و وزن‌ها، گزارش فنی و منابع مرتبط با آموزش RL را نیز منتشر کرده است.