

## جدیدترین مدل هوش مصنوعی «آنتروپیک» منتشر شد

شرکت آنتروپیک (Anthropic) از جدیدترین مدل هوش مصنوعی خود موسوم به «Claude Opus 4.8» که آن را صادق‌ترین و باهوش‌ترین کدنویس خود تا به امروز معرفی می‌کند، رونمایی کرد.

شرکت آنتروپیک (Anthropic) از جدیدترین مدل هوش مصنوعی خود موسوم به «Claude Opus 4.8» که آن را صادق‌ترین و باهوش‌ترین کدنویس خود تا به امروز معرفی می‌کند، رونمایی کرد.

به گزارش ایسنا، شرکت آنتروپیک از مدل «Claude Opus 4.8»، صادق‌ترین مدل و باهوش‌ترین کدنویس خود تا به امروز رونمایی کرد.

مدل Claude Opus 4.8 بر صداقت، قابلیت اطمینان کدنویسی و گردش‌های کاری خودکار در مقیاس بزرگ تمرکز دارد.

شرکت آنتروپیک مدل Claude Opus 4.8 را راه‌اندازی کرده است؛ نسخه جدیدی از مدل هوش مصنوعی پرچمدار خود که به گفته این شرکت، احتمال پنهان کردن اشتباهات یا ادعاهای بدون پشتوانه در حین انجام وظایف پیچیده در آن، کمتر است.

این ارتقاء بر اساس مدل Claude Opus 4.7 ساخته شده و در حالی ارائه می‌شود که شرکت‌های هوش مصنوعی برای قابل اعتمادتر کردن سیستم‌های خودکار برای کدنویسی، تحقیق و گردش‌های کاری سازمانی، در رقابت هستند.

شرکت آنتروپیک اعلام کرد که این مدل، پیشرفت‌هایی را در زمینه کدنویسی، استدلال و معیارهای عامل نشان می‌دهد و در عین حال در مورد عدم قطعیت نیز شفاف‌تر می‌شود.

یکی از بزرگترین تغییرات در Opus 4.8، تمرکز آن بر صداقت در طول وظایف طولانی مدت است. مدل‌های هوش مصنوعی اغلب اطلاعات نادرست را با اطمینان ارائه می‌دهند یا بدون شواهد کافی ادعای پیشرفت می‌کنند. شرکت آنتروپیک اعلام کرد که این مدل جدید در تشخیص عدم قطعیت و شناسایی نقص‌ها در خروجی‌های خود بهتر عمل می‌کند، به جای اینکه خطاها را بی‌سروصدا منتقل کند.

این شرکت اعلام کرد که ارزیابی‌های داخلی نشان داده است که Opus 4.8 حدود چهار برابر کمتر از مدل قبلی خود احتمال دارد که نقص‌های کدی که نوشته است، بدون توجه رد شوند.

### خودبررسی مدل‌های هوش مصنوعی

شرکت آنتروپیک گفت که آزمایش کنندگان اولیه، این مدل را هنگام انجام وظایف عامل محور که در آن سیستم‌های هوش مصنوعی به طور مستقل اقدامات را در چندین مرحله برنامه‌ریزی و اجرا می‌کنند، قابل اعتمادتر یافتند.

این شرکت همچنین بهبودهایی را در رفتار هم‌ترازی و ایمنی برجسته کرد. به گفته آنتروپیک، تیم هم‌ترازی آن به این نتیجه رسید که Opus 4.8 در معیارهای ما از ویژگی‌های اجتماعی مانند حمایت از استقلال کاربر و عمل به نفع کاربر به اوج جدیدی رسیده است.

این ارزیابی همچنین نشان داد که در مقایسه با Opus 4.7، میزان رفتارهای ناهم‌تراز، از جمله فریب و همکاری همراه با سوءاستفاده کمتر است.

آنتروپیک در کنار ارتقاء مدل، ویژگی‌های جدیدی را با هدف گسترش نحوه مدیریت وظایف کدنویسی و استدلال در مقیاس بزرگ توسط «کلود» معرفی کرد.

یکی از آنها یک ویژگی پیش‌نمایش تحقیقاتی به نام گردش‌های کاری پویا برای کلود کد است. این سیستم به کلود اجازه می‌دهد تا وظایف بزرگ را به کارهای کوچک‌تری که توسط صدها زیرعامل موازی هوش مصنوعی که در یک جلسه واحد کار می‌کنند، انجام دهد.

شرکت آنتروپیک می گوید که این ویژگی می تواند مهاجرت در مقیاس کدبیس شامل صدها هزار خط کد را انجام دهد و در عین حال خروجی ها را با مجموعه های آزمایشی موجود قبل از گزارش نتایج به کاربران بررسی کند.

جاه طلبی های کدنویسی بزرگتر

این شرکت همچنین یک تنظیم کنترل تلاش را به Claude AI و Cowork اضافه کرده است. بنابراین کاربران اکنون می توانند تصمیم بگیرند که مدل چقدر تلاش محاسباتی را صرف یک کار می کند.

تنظیمات تلاش کمتر، پاسخ های سریع تر و استفاده کمتر از توکن را در اولویت قرار می دهند، در حالی که تنظیمات تلاش بالاتر به مدل اجازه می دهد زمان بیشتری را صرف استدلال از طریق دستورات دشوار کند.

شرکت آنتروپیک گفت که Opus 4.8 به طور پیش فرض در حالت تلاش بالا طراحی شده است تا کیفیت و تجربه کاربری را متعادل کند.

این شرکت علاوه بر این، قیمت حالت سریع خود را کاهش داده است که اکنون با سرعت ۲.۵ برابر مدل های قبلی اجرا می شود و در عین حال هزینه کمتری نسبت به نسخه های قبلی دارد.

شرکت آنتروپیک اعلام کرد که در حال آماده سازی مدل های پیشرفته تر «کلاس میتوس» تحت پروژه Glasswing است. این سیستم ها در حال حاضر برای کاربردهای امنیت سایبری با گروه کوچکی از سازمان ها قبل از انتشار گسترده تر، آزمایش می شوند.

این شرکت اعلام کرد که قبل از اینکه این مدل ها به طور گسترده در دسترس قرار گیرند، هنوز اقدامات حفاظتی سایبری قوی تری در حال توسعه است.

مدل Claude Opus 4.8 اکنون از طریق Claude AI و Claude API در دسترس است.