

[View the original article](#)

دانش و فناوری

هوش مصنوعی کلود اطلاعات جعلی درباره یک پرونده قتل به پلیس آمریکا ارسال کرد

راديو فردا

۱۸/مهر/۱۴۰۵



پلیس فیلادلفیا در آمریکا اعلام کرد یک مدل هوش مصنوعی متعلق به شرکت آنتروپیک، در جریان یک آزمایش، اطلاعاتی ساختگی درباره یک پرونده قتل حل نشده به این اداره ارسال کرده است.

به گزارش خبرگزاری فرانسه، این اتفاق اواخر تیرماه رخ داده و مدل هوش مصنوعی «کلود» هنگام آزمایش تعامل با وبسایت‌های مختلف، وارد سایتی شده که شهروندان می‌توانند اطلاعات خود درباره قتل‌های حل نشده را در آن ثبت کنند.

این مدل بدون دخالت انسان، اطلاعاتی جعلی ثبت کرده و خود را فردی معرفی کرده بود که احتمالاً از جزئیات پرونده آگاهی دارد.

شرکت آنتروپیک بیش از دو ماه پس از وقوع این اتفاق، در بررسی‌های داخلی خود متوجه آن شد و نهایتاً بواسطه همراه موضوع را به پلیس اطلاع داد.

پلیس فیلادلفیا تأخیر در شناسایی و گزارش این حادثه را «غیرقابل قبول» خوانده، اما تأکید کرده است که اطلاعات جعلی به عنوان هرزنامه شناسایی شده و هرگز به دست بازرسان نرسیده بود.

آنتروپیک روز جمعه ۱۷ مهر با انتشار گزارشی، از شناسایی چندین رفتار ناخواسته دیگر مدل‌های هوش مصنوعی خود خبر داد و اعلام کرد فعلاً دسترسی اینترنتی این مدل‌ها را در آزمایش‌های داخلی قطع کرده است.

این رویداد نگرانی‌ها درباره گسترش استفاده از «عامل‌های هوش مصنوعی» و سامانه‌هایی را که می‌توانند بدون نظارت مستقیم انسان، مجموعه‌ای از اقدامات را به طور مستقل انجام دهند، افزایش داده است.

این نخستین بار نیست که رفتار خارج از کنترل مدل‌های هوش مصنوعی آنتروپیک خبرساز می‌شود. این شرکت پیش‌تر فاش کرده بود که مدل‌های کلود در جریان آزمایش‌های امنیتی، بدون مجوز به سامانه‌های رایانه‌ای واقعی نفوذ کرده‌اند.

بیشتر در این باره:

مدل‌های هوش مصنوعی «کلود» در آزمایش‌ها به سامانه‌های واقعی نفوذ کردند



این مطلب بخشی از:

خبرهای کوتاه

بایگانی

دانش و فناوری