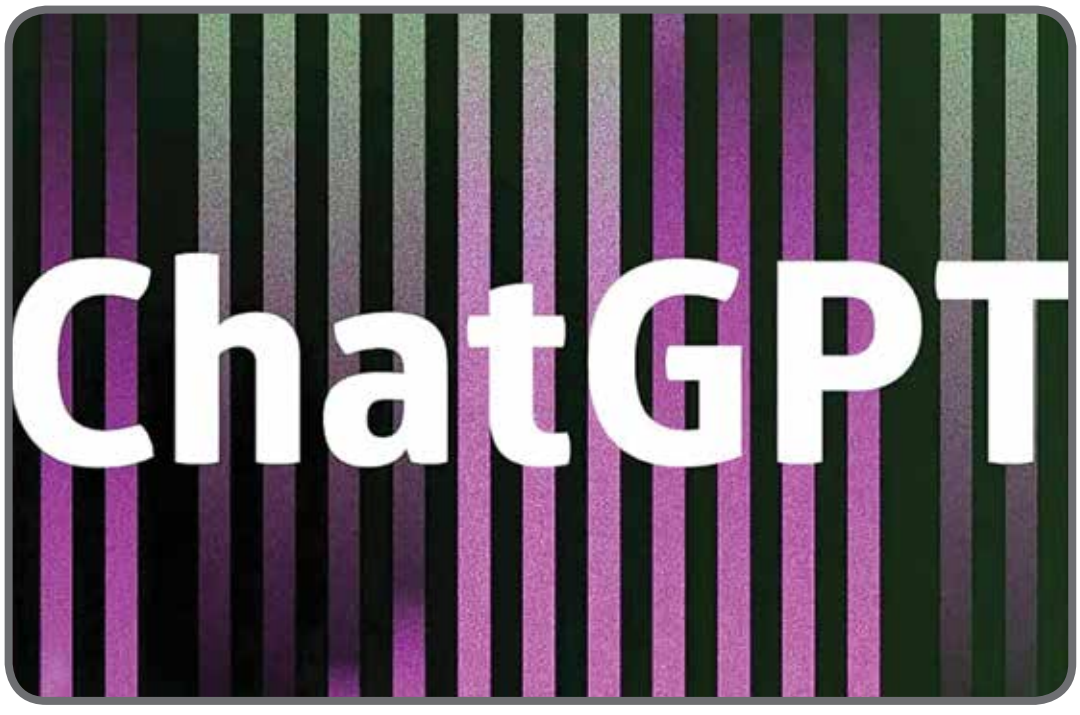
گردانندگان Stack Overflow در پستی می‌گویند: «مشکل اصلی، نرخ بالای نادرست بودن پاسخ‌هایی است که ChatGPT تولید می‌کند، چون معمولا به نظر می‌رسد جواب‌ها خوب و آسان باشد»



این مدل علاوه بر پاسخگویی به سؤالات ساده، کارکردهای بسیاری مانند نوشتن مقاله توصیف هنر با جزئیات زیاد، ایجاد اعلان‌های هنری هوش مصنوعی گفت‌وگوهای فلسفی و حتی کدنویسی دارد

سطح عملکرد ایده‌آل به آموزش و تنظیم دقیق زیادی متکی هستند.

آیا این بدان معناست که هوش مصنوعی جهان را تسخیر کرده است؟ شاید هنوز نه، اما آلتمن از OpenAI فکر می‌کند که هوش به سبک انسانی در هوش مصنوعی اکنون چندان دور نیست. آلتمن در پاسخ به اظهار نظر ماسک در مورد هوش مصنوعی بسیار قوی، در توییتی نوشت: موافقم که به هوش مصنوعی خطرناکی نزدیک هستیم که به عنوان مثال خطر امنیت سایبری بزرگی را به همراه دارد. و فکر می‌کنم می‌توانیم در دهه آینده به AGI واقعی برسیم. بنابراین ما باید خطر آن را نیز بسیار جدی بگیریم.

او همچنین خاطرنشان کرد: مشاهده مردمی که شروع به بحث می‌کنند که آیا سیستم‌های هوش مصنوعی قدرتمند باید به شیوه‌ای که کاربران می‌خواهند یا سازندگان آنها رفتار کنند. این سوال که ما این سیستم‌ها را با چه ارزش‌هایی هماهنگ می‌کنیم از مهم‌ترین بحث‌های جامعه خواهد بود.

سرمایه‌گذاری مایکروسافت

بنا به گزارش‌های اخیر، مایکروسافت برنامه‌های بزرگی برای ویژگی‌های هوش مصنوعی در چند برنامه و خدمات خود دارد. اطلاعاتی که به اشتراک گذاشته شده است که مایکروسافت ChatGPT را در مجموعه آفیس قرار می‌دهد تا به افراد اجازه تولید متن را بدهد. یک گزارش جداگانه توسط همان رسانه بیان کرد که مایکروسافت ChatGPT را در Bing ادغام می‌کند تا پاسخ‌هایی به پرس‌وجوها ایجاد کند.

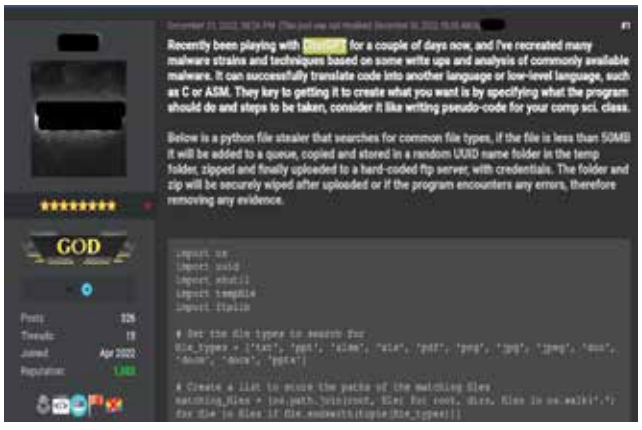
مایکروسافت سال ۲۰۱۹ همکاری خود را با OpenAI اعلام کرد و باعث شد غول فناوری مستقر در ردموند یک میلیارد دلار در OpenAI سرمایه‌گذاری کند. این همکاری همچنین شامل توافق برای استفاده از زیرساخت ابرکامپیوتری هوش مصنوعی Azure مایکروسافت برای آموزش GPT بود که ربات چت معروف ChatGPT را نیرو می‌دهد. به نظر می‌رسد پس از چند سال همکاری، مایکروسافت آماده است تا شاهد بازگشت سرمایه باشد.

بر اساس اطلاعات، مایکروسافت آفیس از کاربران می‌خواهد از دستورات برای تولید متن از طریق هوش مصنوعی پیروی کنند. این تا حدودی شبیه به عملکردی است که آفیس قبلاً دارد، اگرچه احتمالا ویژگی OpenAI-back قدرتمندتر از آنچه در حال حاضر در دسترس است خواهد بود.

در این گزارش آمده که مهندسان و محققان مایکروسافت بیش از یک سال روی ابزارهای هوش مصنوعی برای ایجاد ایمیل و اسناد کار کرده‌اند. یکی از عناصر مهم این ویژگی حفظ حریم خصوصی است. اطلاعات توضیح می‌دهد که مایکروسافت در تلاش است تا اطمینان حاصل کند داده‌های مشتری از طریق ابزار درز نمی‌کند.

هوش مصنوعی در ماه‌های اخیر بر سرفصل‌های دنیای فناوری تسلط داشته است. چت ربات ChatGPT الهام‌بخش حیرت و انتقاد بوده است. ابزار هوش مصنوعی برای ایجاد یک کتاب کودکان استفاده و منجر به اتهام سرقت ادبی شد. VALL-E تازه معرفی شده مایکروسافت که می‌تواند صدای افراد را تولید کند، نگرانی‌های امنیتی را نیز به دنبال داشته است.

در حالی که هوش مصنوعی قدرتمند است و در آینده در بسیاری از برنامه‌ها نقش خواهد داشت، مایکروسافت ممکن است برای اطمینان از استفاده صحیح از آن وقت بیشتری نیاز داشته باشد.



ChatGPT توسط شرکت تحقیقاتی و هوش مصنوعی OpenAI ایجاد شده و آن را در ۳۰ نوامبر ۲۰۲۲ راه‌اندازی کرد

استفاده هکرها

ChatGPT به چشم‌انداز تهدیدات سایبری مدرن نیز چاشنی‌هایی اضافه کرده است، زیرا به سرعت آشکار شد که تولید کد می‌تواند به عوامل تهدید کمتر ماهر کمک کند تا بدون زحمت حملات سایبری انجام دهند. سوال نخست این‌که فقط یک تهدید فرضی است یا عوامل تهدیدی وجود دارد که از فناوری‌های OpenAI برای اهداف مخرب استفاده می‌کند.

تجزیه و تحلیل CPR از چند انجمن اصلی هک زیرزمینی نشان می‌دهد اولین نمونه از مجرمان سایبری وجود دارد که از OpenAI برای توسعه ابزارهای مخرب استفاده می‌کند. همانطور که مشکوک بودیم، برخی موارد به وضوح نشان داد بسیاری از مجرمان سایبری که از OpenAI استفاده می‌کنند، اصلا مهارت توسعه ندارند. گرچه ابزارهایی که در این گزارش ارائه می‌کنیم بسیار ابتدایی است، اما تا زمانی که عوامل تهدید پیچیده‌تر، روشی را که از ابزارهای مبتنی بر هوش مصنوعی برای بد استفاده می‌کنند، بهبود بخشند زمان زیادی نیست.

در ۲۹ دسامبر ۲۰۲۲، موضوعی به نام « ChatGPT - مزایای بدافزار » در یک انجمن هک زیرزمینی محبوب ظاهر شد. ناشر این موضوع فاش کرد که با ChatGPT در حال آزمایش برای بازآفرینی گونه‌های بدافزار و تکنیک‌هایی است که در نشریات تحقیقاتی و نوشته‌های مربوط به بدافزار رایج شرح داده شده‌اند. به عنوان مثال، او کد یک دزد مبتنی بر پایتون را به اشتراک گذاشت که انواع فایل‌های رایج را جست‌وجو می‌کند، آنها را در یک پوشه تصادفی داخل پوشه Temp کپی می‌کند، آنها را ZIP و در یک سرور FTP با کد سخت آپلود می‌کند.

تجزیه و تحلیل ما ادعاهای مجرم سایبری را تأیید می‌کند. این در واقع یک دزد است که ۱۲ نوع فایل رایج (مانند اسناد PDF، MS Office و تصاویر) را در سراسر سیستم جست‌وجو می‌کند. اگر فایل‌های مورد علاقه پیدا شد، بدافزار، فایل‌ها را در فهرستی موقت کپی و فشرده می‌کند و از طریق وب می‌فرستد. شایان ذکر است که این بازیگر زحمت رمزگذاری یا ارسال ایمن فایل‌ها را نداشته است، بنابراین ممکن است فایل‌ها به دست اشخاص ثالث نیز برسد.

نمونه دومی که این بازیگر با استفاده از ChatGPT ایجاد کرد، یک قطعه ساده جاوا است. PuTTY، یک کلاینت بسیار رایج SSH و telnet را دانلود کرده و به طور مخفیانه با استفاده از Powershell روی سیستم اجرا می‌کند. البته می‌توان این اسکریپت را برای دانلود و اجرای هر برنامه‌ای از جمله خانواده بدافزارهای رایج تغییر داد.

مشارکت قبلی این عامل تهدید در تالار گفتمان شامل به اشتراک‌گذاشتن چند اسکریپت مانند اتوماسیون مرحله پس از بهره‌برداری و یک برنامه C++ است. علاوه بر این، او به طور فعال نسخه‌های کرک شده SpyNote، یک بدافزار Android RAT را به اشتراک می‌گذارد. بنابراین به طور کلی، به نظر می‌رسد این فرد یک عامل تهدید فناوری محور است و هدف از پست‌های او این است که به مجرمان سایبری با توانایی فنی کمتر نشان دهد چگونه از ChatGPT برای اهداف مخرب استفاده کنند.