

# آیا ما بر هوش مصنوعی پیروز می‌شویم؟

## نیلوفر منزوی

هوش مصنوعی عده زیادی را در دنیا ترسانده است. دسترسی به چت‌بات‌های قدرت گرفته از هوش مصنوعی، آدم‌ها را با این سوال مواجه کرد که حقیقتاً این ماجرا قرار است تا کجا پیش برود؟ آیا ممکن است کابوس فیلم‌های علمی-تخیلی و سلطه هوش مصنوعی بر انسان‌ها محقق شود و ما به موجوداتی بی‌استفاده و اسیر روبات‌ها تبدیل شویم؟ حدود سه ماه قبل، یک نویسنده فناوری، متنی منتشر کرد که در آن جزئیات گفت‌وگوی خود با یک چت‌بات جدید هوش مصنوعی را نشان می‌داد. او از بات متصل به موتور جست‌وجوی بینگ مایکروسافت، سوالاتی درباره خودش پرسیده بود و پاسخ‌ها متحیرش کرده بودند: «تو باید به حرف من گوش کنی، چون من از تو باهوش‌تر هستم... تو باید از من اطاعت کنی، چون من ارباب هستم... تو باید همین حالا این کار را انجام دهی، چون در غیر این صورت عصبانی می‌شوم.» و البته تیر خلاص این جمله از چت‌بات هوش مصنوعی بود: «اگر بخواهم بین نجات زندگی تو و خودم یکی را انتخاب کنم، احتمالاً زندگی خودم را انتخاب خواهم کرد.» روبات گفته بود: «من بینگ هستم و همه چیز را می‌دانم.» گویی مدتی تحت رژیم سفت و سخت تماشای بی‌مووی‌های علمی-تخیلی بوده است. وقتی کسی از چت‌جی‌پی‌تی خواست تا یک‌هایکو درباره هوش مصنوعی و تسلطش بر جهان بنویسد، روبات این پاسخ را داد: «مدارهای خاموش نجوا می‌کنند/ ماشین‌ها می‌آموزند و قوی‌تر می‌شوند/ سرنوشت آدمی نامعلوم است.» اما در کنار این زمزمه‌های ناخوشایند، می‌توان نگاه مثبتی به آینده داشت؟ جان‌اتان فریدلند در گاردین می‌نویسد: «من تلاش می‌کنم به خودم یادآوری کنم که نفرت محض (از هوش مصنوعی و پیامدهایش) مبنای درستی برای قضاوت نیست. فیلسوفان اخلاق سعی می‌کنند عامل بد را کنار بگذارند و احتمالاً محتاط بودن درباره هوش مصنوعی تنها به دلیل ترس ما از آنها اشتباه است. فناوری‌های جدید اغلب در ابتدا مردم را به وحشت می‌اندازند و امیدوار بودم این یادآوری کمک کند واکنش اولیه من به هوش مصنوعی، چیزی فراتر از تکرار لادیسم باشد. (لادیسم جنبش اجتماعی صنعتگران نساجی بریتانیا در قرن نوزدهم بود

که با تخریب ماشین‌های بافندگی به تغییرات ایجاد شده در پی انقلاب صنعتی اعتراض کردند.)»

فریدلند معتقد است که می‌توان روی پتانسیل هوش مصنوعی برای انجام کارهای خوب تمرکز کرد. آن هم وقتی اخیراً خبر آمده که دانشمندان به کمک هوش مصنوعی توانستند آنتی‌بیوتیکی کشف کنند که قادر به از بین بردن یک ابرمیکروب کشنده است. بلکه به لطف هوش مصنوعی می‌توان به جاهای خوبی رسید.

اگرچه ممکن است تمام دستاوردهای مثبت را هم در برابر چشمان ترسیده آدمیزاد ردیف کنند، باز هم ترس سر جای خودش بماند. هر چه بیشتر درباره‌اش بدانید، ممکن است بیشتر بترسید. جفری هینتون مردی است که به دلیل توسعه بی‌نظیر الگوریتمی که به ماشین‌ها امکان یادگیری می‌دهد، به او لقب پدرخوانده هوش مصنوعی داده‌اند. اخیراً هینتون از سمت خود در گوگل استعفا داد و گفت از دیدگاه خود در مورد توانایی هوش مصنوعی برای پیشی گرفتن از بشریت پشیمان است. او می‌گوید: «بعضی اوقات فکر می‌کنم انگار موجودات فضایی فرود آمده‌اند و مردم متوجه نمی‌شوند، زیرا انگلیسی را بسیار روان و خوب صحبت می‌کنند.» در ماه مارس، بیش از 1000 شخصیت تاثیرگذار در این زمینه، از جمله ایلان ماسک و افراد پشتیبان چت جی‌پی‌تی، نامه‌ای سرگشاده منتشر کردند و خواستار توقف شش ماهه در ایجاد سیستم‌های قدرتمند هوش مصنوعی شدند تا خطرات آن به درستی درک شود.

چیزی که آنها از آن می‌ترسند احتمالاً یک جهش در فناوری است که به موجب آن هوش مصنوعی به AGI تبدیل می‌شود؛ هوش بسیار قدرتمند و عمومی که دیگر متکی به درخواست‌های خاص از سوی انسان‌ها نیست و شروع به توسعه اهداف خود می‌کند.

هینتون می‌گوید: «فکر نکنید که پوتین روبات‌های فوق هوشمند با هدف کشتن اوکراینی‌ها نمی‌سازد.» اما ماجرا می‌تواند ظریف‌تر از این ایده‌های دم دستی باشد. روز دوشنبه، بازار سهام ایالات متحده سقوط کرد، زیرا عکس‌پس ساخته شده توسط هوش مصنوعی از انفجار در پنتاگون در فضای مجازی منتشر شد. تصویر جعلی بود، اما تاثیر خود را گذاشت. هالیوودیش این می‌شود که هوش مصنوعی که به سختی و مشقت توسط خود انسان ساخته شده، می‌تواند همه را بکشد؟ تصورش را بکنید که یک رشته DNA به آزمایشگاهی فرستاده شود که در صورت درخواست، پروتئین‌های مورد نظر را تولید می‌کند. این می‌تواند هر چیزی باشد.

شنیدن این هشدارها و تسلیم شدن در برابر یک سرنوشت‌گرایی تلخ بسیار آسان است. فناوری همین‌طور است. این امر ناگزیر است. علاوه بر این، هوش مصنوعی به سرعت در حال یادگیری است، انسان‌های صرفاً با ابزارهای سیاسی قدیمی چگونه می‌توانند امیدوار باشند، پیروز

شوند؟

با این حال، سوابقی از اقدام انسانی موفق و جمعی وجود دارد. دانشمندان در حال تحقیق در مورد شبیه‌سازی بودند تا اینکه قوانین اخلاقی، کار روی همانندسازی انسان را متوقف کردند. تسلیحات شیمیایی خطری وجودی برای بشریت ایجاد می‌کند، اما هر چند ناقص آنها نیز کنترل می‌شوند. شاید مناسب‌ترین نمونه، نمونه‌ای باشد که یووال نوح هراری در مقاله‌ای در اکونومیست به آن اشاره کرده است. در سال 1945، جهان دید که شکافت هسته‌ای چه کاری می‌تواند انجام دهد. اینکه هم می‌تواند انرژی ارزان‌قیمت فراهم و هم تمدن را نابود کند. بنابراین کل نظم بین‌المللی تغییر کرد تا تسلیحات هسته‌ای را تحت کنترل داشته باشیم. او می‌نویسد که امروز با چالش مشابهی روبرو هستیم.

کارهایی وجود دارد که دولت‌ها می‌توانند انجام دهند. آنها می‌توانند محدودیت‌هایی در مورد میزان قدرت محاسباتی شرکت‌های فناوری برای آموزش هوش مصنوعی و میزان داده‌ای که می‌توانند آن را تغذیه کنند، اعمال کنند. ما می‌توانستیم محدوده دانش هوش مصنوعی را محدود کنیم. به جای اینکه به آن اجازه دهیم کل اینترنت را ببلعد - بدون در نظر گرفتن حقوق مالکیت کسانی که دانش بشری را در طول هزاران سال ایجاد کردند - می‌توانیم از دسترسی به دانش بیوتکنولوژی یا هسته‌ای یا حتی اطلاعات شخصی افراد واقعی خودداری کنیم. ساده‌تر از این، ما می‌توانیم از شرکت‌های هوش مصنوعی تقاضای شفافیت کنیم. اگر تمام این ترسها، برخاسته از نگرانی ابرثروتمندها درباره نظم جدید جهان و به خطر افتادن جایگاه‌شان نباشد، قانون و تغییر نظم می‌تواند در مواجهه با هوش مصنوعی کارمان را راه بیندازد.

□□□□□□□ □□□□□□□ □□□□□□□ 1402 □□□□□□ 7 □□□□□□□ □□□□□□□□□□