

چطور در مواجهه با هوش مصنوعی دچار توهم تسلط و اعتماد به نفس کاذب نشویم

علی نجفی، ۱۶ شهریور ۱۴۰۵

با آسان شدن دسترسی به مدل‌های زبانی مثل چت‌جی‌پی‌تی و دیگران که به اختصار آنها را هوش مصنوعی می‌نامیم، لازم شده است تا عمیق‌تر به نحوه‌ی تعامل با آنها توجه کنیم. آیا تا به حال شده است که با چندین سوال پی‌درپی در موضوعی خاص و دیدن پاسخ‌های تمیز، جمع‌وجور و فرمول‌های کادربندی شده‌ی هوش مصنوعی که با تبحر کافی و با سبک لاتک حروفچینی شده به این نتیجه برسید که چقدر موجود خوبی است و جواب سوالات شما را داده و شما با رضایت و اعتماد به نفس، خودتان را متخصص آن موضوع بدانید؟

آیا تا به حال مقایسه‌ای انجام داده‌اید بین نتایج این دو حالت: با سوالی در دست به سراغ یکی از معلمان با تجربه می‌رویم و دفعه‌ی بعدی با همان سوال به سراغ هوش مصنوعی می‌رویم؟ یک موضوع مهم که مورد بحث است، توهم چیرگی یا توهم تسلط (illusion of competence) و اعتماد به نفس کاذبی است که گاهی در مواجهه با هوش مصنوعی به دست می‌آید. سوالات را در ۲ دسته قرار دهیم،

نوع یکم: سوالات کاملاً مشخص و با صورت‌بندی دقیق. مثال: معادله‌ی نوسانگر هارمونیک و ریاضی یک بعدی کلاسیکی با شرط اولیه داده شده را با چه ترفندی حل کنم، جواب هم معمولاً شبیه این است: با روش تبدیل لاپلاس که بصورت زیر است ...

نوع دوم: سوالاتی‌اند که هنوز صورت پخته‌ای ندارند و اصلاً خود سوال هم نامشخص است. مثلاً: طیف اتم هیدروژن را با چه روشی محاسبه کنم؟ جنبه‌های زیادی در این سوال نامشخص است: معلوم نیست اثرهای نسبیتی مورد نظر سوال کننده هست یا نه؟ با چه تقریبی می‌خواهد طیف اتم را به دست آورد؟ اتم هیدروژن روی زمین است یا خورشید؟ فضا چند بعدی است و بیشتر سوالات ما از نوع دوم است.

با سوالی از نوع دوم ولی تصور اینکه نوع اول است به سراغ هوش مصنوعی برویم به احتمال زیاد دچار توهم چیره‌گی می‌شویم. با مجموعه‌ای از نوشته‌ها و روابط ریاضی مواجه می‌شویم که در دید اول به نظر منطقی، جامع، منسجم، روان و مهیج می‌آیند ولی با کمی عمیق شدن و تحمل زحمت باز تولید برخی از آنها، معلوم می‌شود که در واقع بیشترشان اشتباه هستند. حتی اگر درست هم باشند، دیدن روابط و خواندن متون و مفتون شیوایی و انسجام متن شدن، لزوماً به معنی درک و یادگیری آنها نیست. یک معیار ساده این است: آیا می‌توانیم همان استدلال را با جمله‌بندی‌های خودمان باز تولید و به شخص دیگری انتقال دهیم. آیا می‌توانیم یک انشای ۱ صفحه‌ای در مورد آن موضوع بنویسیم؟

پس چکار کنیم: در حل هر مساله‌ای قاعدتاً این مراحل طی می‌شود: جنبه‌ها و ایده‌های مختلف سنجیده می‌شوند. فرضیات دقیقاً مشخص می‌شوند. قیده‌های لازم فیکس می‌شوند. تمام شفاف سازی‌ها صورت می‌گیرد تا صورت مساله شکل واقعی به خود بگیرد. صورت مساله که واضح شد، آن را تفکیک می‌کنیم به قدم‌های کوچک‌تر. هر قدم را جداگانه بررسی می‌کنیم. در نهایت جمع بندی می‌کنیم. نمی‌توانیم از

هوش مصنوعی بخواهیم همه‌ی این مراحل را خود و به یکباره طی کند. از طرف دیگر در طی کردن همه‌ی این مراحل می‌توان روی کمک هوش مصنوعی حساب کرد. فقط باید تفکیک درستی از نقش خودمان و نقش هوش مصنوعی در هر مرحله داشته باشیم. با داشتن درک درست از وظایف و نقش خودمان، و نادیده نگرفتن و عدم انکار نقش خودمان، می‌توانیم جواب‌های منطقی از هوش مصنوعی بگیریم و دچار توهم چیرگی هم نشویم. اگر توانستیم با هوش مصنوعی وارد یک بحث عمیق علمی در زمینه‌ی تمام مراحل شویم، احتمالاً راهمان درست است ولی اگر به سرعت جواب‌های تمیزی تولید شد، نمی‌توان زیاد امیدوار بود. معلم با تجربه فوراً سراغ پاسخ نمی‌رود و در عوض به شناسایی و تعریف صورت مساله وزن زیادی می‌دهد. تا تمام جنبه‌ها شناسایی نشود سراغ جواب نمی‌رود.

یادمان باشد: مرجع نهایی تشخیص درستی و حقیقت قاعدتاً نباید هوش مصنوعی شود، فقط آن را بصورت یک ابزار کمکی، مشورتی و جستجوگر در نظر بگیریم. محاسبه را کنار نگذاریم. قدم‌های ریاضی را خودمان بیازماییم. فریب شیوایی و روانی متون و روابط ریاضی را نخوریم. با هر پاسخی از جانب هوش مصنوعی، با نگاهی سختگیرانه و منتقدانه برخورد کنیم.