



هوش مصنوعی و آینده‌ی میدان‌های عمومی دیجیتال

نویسندگان: گلدبرگ‌بث، دیانا و سایرین | مجله هوش جمعی | ۲۰۲۶



چکیده:

دو پیشرفت فناورانه چشمگیر در دهه‌های اخیر، عرصه عمومی را دستخوش تغییر کرده‌اند: نخست، ظهور اینترنت و دوم، معرفی اخیر مدل‌های زبانی بزرگ (LLMها). مدل‌های زبانی بزرگ فرصت‌هایی برای بازاندیشی و بازطراحی فضاهای آنلاین غیرمتمرکزتر و مشارکت‌محورتر فراهم می‌کنند که می‌توان از آن‌ها برای تسهیل گفت‌وگوهای تعاملی و مبتنی بر بحث و استدلال در مقیاس گسترده استفاده کرد؛ با این حال، مدل‌های زبانی بزرگ در صورت نبود حکمرانی مناسب می‌توانند شکاف‌های اجتماعی را نیز تشدید کنند. در اینجا، چهار کاربرد مدل‌های زبانی بزرگ برای بهبود عرصه‌های عمومی دیجیتال را بررسی می‌کنیم: سامانه‌های گفت‌وگوی جمعی، سامانه‌های پل‌ساز، نظارت و تعدیل اجتماعی، و سامانه‌های اثبات انسان‌بودن. با تکیه بر دیدگاه‌های بیش از ۷۰ متخصص جامعه مدنی و فناوری، استدلال می‌کنیم که مدل‌های زبانی بزرگ هم فرصت‌های امیدبخشی برای تغییر پارادایم گفت‌وگو در مقیاس گسترده فراهم می‌کنند و هم خطرات متمایزی برای عرصه‌های عمومی دیجیتال به همراه دارند. از آنجا که افراد از طریق تعاملات خود در عرصه‌های عمومی شکل می‌گیرند، فرصت مهمی برای تقویت ظرفیت‌های فردی و انسجام جمعی از طریق دگرگون‌سازی این فضاهای عمومی دیجیتال وجود دارد. در پایان، دستورکاری برای پژوهش‌ها و سرمایه‌گذاری‌های آینده در حوزه هوش مصنوعی ارائه می‌کنیم که به تقویت عرصه‌های عمومی دیجیتال و محافظت در برابر سوءاستفاده‌های احتمالی از هوش مصنوعی کمک خواهد کرد.

کلمات کلیدی: دموکراسی، هوش مصنوعی، فناوری‌های گفت‌وگو محور، انسجام اجتماعی، تعدیل محتوا



خلاصه مبسوط مقاله:

مقاله «هوش مصنوعی و آینده میدان‌های عمومی دیجیتال» به بررسی تأثیر دو تحول فناورانه مهم، یعنی ظهور اینترنت و توسعه مدل‌های زبانی بزرگ (LLMها)، بر میدان عمومی و فرایندهای دموکراتیک می‌پردازد. نویسندگان معتقدند اینترنت اگرچه امکان مشارکت گسترده‌تر، دسترسی به اطلاعات و شکل‌گیری گفت‌وگوی عمومی را فراهم کرده است، اما میدان‌های عمومی دیجیتال موجود همچنان با مشکلاتی مانند نابرابری در دسترسی، آزار و حذف گروه‌های حاشیه‌ای، نظارت و نقض حریم خصوصی، قطبی‌شدن، اتاق‌های پژواک، تبلیغات دستکاری‌گرانه و فعالیت گروه‌های افراطی مواجه‌اند. از نظر مقاله، پیشرفت LLMها فرصتی برای بازطراحی این فضاها در جهت مشارکت بیشتر، گفت‌وگوی سازنده و انسجام اجتماعی ایجاد کرده است، اما این فناوری‌ها در صورت نبود حکمرانی مناسب می‌توانند شکاف‌های اجتماعی و مشکلات موجود را نیز تشدید کنند. بر همین اساس، مقاله چهار حوزه اصلی کاربرد هوش مصنوعی را بررسی می‌کند: سامانه‌های گفت‌وگوی جمعی، سامانه‌های پل‌ساز، تعدیل مبتنی بر جامعه و سامانه‌های اثبات انسان‌بودن.

مقاله با تکیه بر مفهوم «میدان عمومی» توضیح می‌دهد که اینترنت از ابتدا به‌عنوان فضایی برای گفت‌وگوی عمومی، شکل‌گیری اجتماعات و مشارکت دموکراتیک تصور شده بود و با کاهش موانع تولید و گردش اطلاعات، امکان مشارکت گروه‌های بیشتری را فراهم کرد. با این حال، میدان عمومی دیجیتال هنوز با آرمان گفت‌وگوی آزاد، برابر و مبتنی بر درک متقابل فاصله دارد. نابرابری در دسترسی به فناوری، آزار گروه‌های حاشیه‌ای، نظارت گسترده، نقض حریم خصوصی و خودسانسوری از یک سو، و سازوکارهای پلتفرمی مبتنی بر بیشینه‌سازی تعامل، اتاق‌های پژواک، تبلیغات سیاسی هدفمند و فعالیت گروه‌های افراطی از سوی دیگر، کیفیت گفت‌وگوی عمومی را تضعیف می‌کنند. این فرایندها می‌توانند جامعه را قطبی‌تر از واقعیت نشان دهند، اعتماد میان گروه‌ها و میان شهروندان و نهادهای مدنی را کاهش دهند و مشارکت معنادار در زندگی مدنی را محدود کنند. از این منظر، مقاله استدلال می‌کند که پیشرفت LLMها فرصتی برای بازطراحی میدان عمومی دیجیتال و نزدیک کردن آن به الگویی مشارکت‌محور، فراگیر و مبتنی بر گفت‌وگوی سازنده است.

در بخش نخست، مقاله سامانه‌های گفت‌وگوی جمعی (CDS) را بررسی می‌کند که میان نظرسنجی و گروه کانونی قرار می‌گیرند و امکان مشارکت تعداد زیادی از افراد، بیان دیدگاه‌ها با زبان خودشان، تعامل تدریجی با نظرات دیگران و ترکیب دیدگاه‌های مشارکت‌کنندگان را فراهم می‌کنند. این سامانه‌ها می‌توانند در تصمیم‌گیری‌های رسمی و غیررسمی، از جمله سیاست‌گذاری، مجامع شهروندی و شکل‌دهی دستورکار عمومی، به شناسایی نقاط اشتراک و ارائه راه‌حل‌های جدید کمک کنند. با وجود این، استفاده از CDSها همچنان به نیروی انسانی، تخصص آماری، هزینه و ظرفیت نهادی قابل توجهی نیاز دارد. مقاله نشان می‌دهد LLMها می‌توانند در آموزش تسهیل‌گران، تولید محتوای آموزشی، ترجمه هم‌زمان، طراحی پرسش‌ها و سناریوها، بررسی واقعیت، پیش‌بینی رأی، خلاصه‌سازی و تجسم دیدگاه‌های جمعی و ایجاد بسترهای یکپارچه برای فرایندهای گفت‌وگو به کار گرفته شوند. با این حال، خطرهایی مانند غلبه دیدگاه اکثریت بر اقلیت، تحریف نظرات، تولید محتوای ساختگی، کاهش عاملیت شهروندان و آسیب به مشروعیت فرایندهای دموکراتیک وجود دارد؛ بنابراین LLMها باید مکمل فرایند گفت‌وگو باشند، نه جایگزین مشارکت انسانی.





دومین حوزه، «سامانه‌های پل‌ساز» است که هدف آن ایجاد ارتباط میان افراد دارای دیدگاه‌ها، ارزش‌ها و هویت‌های متفاوت و تقویت اعتماد و نقاط اشتراک میان آنان است. نویسندگان معتقدند الگوریتم‌های رایج توصیه محتوا اغلب به دلیل تمرکز بر تعامل و جذب توجه، محتوای قطبی‌کننده را تقویت می‌کنند و می‌توانند به اتاق‌های پژواک، تشدید اختلافات و حتی رادیکال‌شدن منجر شوند. در مقابل، الگوریتم‌های پل‌ساز می‌توانند محتوایی را برجسته کنند که مورد استقبال گروه‌های متنوع قرار گرفته یا ویژگی‌هایی مانند سازندگی، کنجکاوی، همدلی، بیان تجربه شخصی و ارائه راه‌حل دارد. نمونه‌هایی مانند Community Notes و برخی آزمایش‌های پلتفرم‌ها نشان می‌دهند که استفاده از «تعامل متنوع» می‌تواند کیفیت محتوا را افزایش دهد و قطبی‌شدن را کاهش دهد. با این حال، مقاله تأکید می‌کند که افزایش صرفِ مواجهه با دیدگاه مخالف الزاماً به کاهش قطبی‌شدن منجر نمی‌شود و محتوای خصمانه حتی می‌تواند آن را افزایش دهد. همچنین، سامانه‌های پل‌ساز ممکن است به دلیل شفاف نبودن، دستکاری الگوریتم، تقویت محتوای سطحی و نگرانی‌های حریم خصوصی با مشکل مشروعیت مواجه شوند؛ از این‌رو پژوهش بیشتر درباره سنجش اثربخشی، اعتماد، شفافیت و طراحی این سامانه‌ها ضروری است.

سومین حوزه، تعدیل محتوای مبتنی بر جامعه است که در آن مدیریت و تعیین هنجارهای گفت‌وگو تا حدی به خود اجتماعات آنلاین واگذار می‌شود. مقاله میان تعدیل تجاری در سطح پلتفرم و تعدیل جامعه‌محور توسط اعضای اجتماعات تمایز می‌گذارد و نشان می‌دهد که هنجارهای قابل قبول در هر جامعه می‌توانند متناسب با زمینه متفاوت باشند. با این حال، تعدیل‌کنندگان داوطلب با مشکلاتی مانند حجم بالای کار، فرسودگی، فشار عاطفی، کمبود آموزش و ابزار، تعارض با کاربران و نبود سازوکارهای مناسب اعتراض مواجه‌اند. LLMها می‌توانند از طریق تحلیل و خلاصه‌سازی تعاملات، شبیه‌سازی بحث‌ها، تدوین دستورالعمل‌های جامعه، ارائه بازخورد پیش از انتشار، یادآوری قواعد، کاهش تنش، دسته‌بندی گزارش‌ها و پشتیبانی از فرایندهای اعتراض به تعدیل‌کنندگان کمک کنند. در عین حال، استفاده از هوش مصنوعی می‌تواند سوگیری‌های زبانی و حذف صداهای اقلیت را ایجاد کند و نمی‌تواند جایگزین انسان در مدیریت تعارضات رابطه‌ای و بحران‌های حساس شود. بنابراین مقاله بر حمایت از تعدیل‌کنندگان انسانی، جبران کار آنان، مشارکت جامعه در تعیین هنجارها و حفظ انسان در چرخه تصمیم‌گیری تأکید می‌کند.

چهارمین حوزه، سامانه‌های «اثبات انسان‌بودن» است که با هدف تشخیص کاربران انسانی از حساب‌ها و عوامل خودکار اهمیت بیشتری پیدا کرده‌اند؛ زیرا توسعه هوش مصنوعی امکان تولید انبوه محتوا، ایجاد شخصیت‌های جعلی و فعالیت گسترده‌تر ربات‌ها را فراهم کرده است. مقاله توضیح می‌دهد که این سامانه‌ها می‌توانند از طریق اتصال حساب‌های آنلاین به هویت‌های واقعی یا استفاده از اعتبارنامه‌های شخص‌بودگی و روش‌هایی مانند اثبات‌های دانش صفر، انسانی‌بودن کاربران را بدون افشای کامل هویت آنان تأیید کنند. با این حال، میان امنیت در برابر دستکاری و اصولی مانند حریم خصوصی، ناشناس‌بودن، استقلال فردی و دسترسی برابر تعارض‌هایی وجود دارد و هیچ روش اثبات انسان‌بودنی کاملاً مصون از خطا نیست. مقاله برای طراحی مسئولانه این سامانه‌ها بر پاسخگویی، اثربخشی، فراگیری، قابلیت همکاری، حفظ حریم خصوصی، تناسب، امنیت و حاکمیت فردی تأکید می‌کند. همچنین پیشنهاد می‌کند پژوهش‌های آینده بر روش‌های تأیید ناشناس، اعتبارنامه‌های شخص‌بودگی، اثبات‌های دانش صفر و استانداردهای قابل حمل و سازگار با تحولات سریع هوش مصنوعی تمرکز کنند.

در نتیجه، مقاله استدلال می‌کند که LLMها می‌توانند به بازسازی میدان عمومی دیجیتال و تقویت فرایندهای دموکراتیک کمک کنند، مشروط بر اینکه به‌عنوان ابزارهای تقویت‌کننده ظرفیت انسانی و جمعی، نه جایگزین انسان، به کار گرفته شوند. سامانه‌های

گفت‌وگوی جمعی می‌توانند مشارکت و شنیده‌شدن دیدگاه‌های شهروندان را افزایش دهند، سامانه‌های پل‌ساز می‌توانند انگیزه‌های قطبی‌کننده را کاهش دهند، تعدیل جامعه‌محور می‌تواند مشروعیت و پاسخگویی مدیریت اجتماعات را افزایش دهد و سامانه‌های اثبات انسان‌بودن می‌توانند از دستکاری توسط عوامل غیرواقعی جلوگیری کنند. تحقق این ظرفیت به سرمایه‌گذاری میان‌رشته‌ای، ابزارهای متن‌باز، مشارکت جامعه مدنی، همکاری صنعت و دانشگاه و تنظیم‌گری مسئولانه نیاز دارد. مقاله همچنین بر اهمیت ارزیابی تجربی، شفافیت، اعتماد عمومی و مشارکت کاربران در طراحی، اجرا و ارزیابی این فناوری‌ها تأکید می‌کند. در نهایت، آینده میدان عمومی دیجیتال نه صرفاً به پیشرفت فناوری، بلکه به چگونگی حکمرانی آن، نحوه توزیع قدرت و میزان قرارگرفتن «عموم مردم» در مرکز طراحی و تصمیم‌گیری وابسته است.

نکات کلیدی:

- موضوع مطالعه: نقش مدل‌های زبانی بزرگ (LLM) ها (در بازطراحی و تقویت میدان‌های عمومی دیجیتال و فرایندهای گفت‌وگوی دموکراتیک
- مفهوم مرکزی: میدان عمومی دیجیتال به مثابه فضایی برای مشارکت، گفت‌وگو، شکل‌گیری ظرفیت‌های فردی و تقویت انسجام جمعی
- چارچوب تحلیلی: بررسی چهار کاربرد اصلی LLM ها شامل سامانه‌های گفت‌وگوی جمعی، سامانه‌های پل‌ساز، تعدیل محتوای مبتنی بر جامعه و سامانه‌های اثبات انسان‌بودن
- یافته کلیدی LLM: ها می‌توانند مشارکت گسترده‌تر، گفت‌وگوی سازنده، کاهش قطبی‌شدن، بهبود تعدیل محتوا و مقابله با فعالیت عوامل غیرانسانی را در میدان‌های عمومی دیجیتال تسهیل کنند
- استدلال اصلی: ظرفیت LLM ها برای تقویت میدان عمومی دیجیتال تنها در صورتی محقق می‌شود که این فناوری‌ها با حکمرانی مناسب، شفافیت، مشارکت انسانی، حفظ حریم خصوصی و سازوکارهای پاسخگویی همراه باشند و به‌عنوان مکمل، نه جایگزین، کنش و داوری انسانی به کار گرفته شوند



اندیشکده فرهنگ و توسعه

باشگاه مقالات فرتو