

در بخش سوم از سری محسنات و خطرات هوش مصنوعی به خطرات احتمالی هوش مصنوعی پرداخته میشود.

فرصت ها و تهدید ها:

با وجود این نگرانی‌ها، مدل‌ها و فناوری‌های هوش مصنوعی نیز فرصت‌های مهمی را ارائه می‌دهند، به عنوان مثال بهبود کارایی و اثربخشی در زمینه‌هایی مانند مراقبت‌های بهداشتی، اجرای قانون و بسیاری موارد دیگر. چالش این است که این فناوری‌ها را به گونه‌ای طراحی و تنظیم کنیم که مزایا را به حداکثر برساند و معایب را به حداقل برساند. بحث در مورد هوش مصنوعی فقط در مورد پیشرفت‌های تکنولوژیکی نیست، بلکه در مورد مسائل اخلاقی، اجتماعی و سیاسی عمیق است. توسعه بیشتر و ادغام هوش مصنوعی در جامعه نیازمند رویکردی متعادل است که هم پتانسیل عظیم و هم خطرات را در نظر می‌گیرد.

بسیار مهم است که چالش‌های مرتبط با هوش مصنوعی را بشناسیم و به آنها پاسخ دهیم تا آینده‌ای را تضمین کنیم که در آن از فناوری‌های هوش مصنوعی به نفع همه استفاده شود. توسعه و استفاده از هوش مصنوعی باید با **درک عمیق مسائل اخلاقی، خطرات امنیتی** و احتمال به نابودی بشر هدایت شود. با تعادل مناسب بین شناسایی فرصت‌ها و مدیریت ریسک‌ها، هوش مصنوعی می‌تواند نیروی محرکه‌ای برای تغییرات مثبت باشد.

استفاده‌های احتمالی از هوش مصنوعی نیز برای شرکت‌ها اهمیت فزاینده‌ای پیدا می‌کند، زیرا آنها مزایای مختلفی را به همراه دارند. یک استراتژی هوش مصنوعی سنجیده و جامع برای تضمین مزیت رقابتی پایدار بسیار مهم است. شرکت‌هایی که از یک برنامه مشخص پیروی می‌کنند از هوش مصنوعی به طور موثرتر و موفقیت‌آمیزی استفاده می‌کنند. بنابراین برنامه‌ریزی و استراتژی برای موفقیت پروژه‌های هوش مصنوعی ضروری است.

مواردی از معایب هوش مصنوعی:

آیا هوش مصنوعی خطرناک است؟

جامعه فناوری مدت‌هاست در مورد تهدیدات ناشی از هوش مصنوعی بحث کرده است. اتوماسیون مشاغل، انتشار اخبار جعلی و مسابقه تسلیحاتی خطرناک تسلیحات مجهز به هوش مصنوعی به عنوان برخی از بزرگترین خطرات ناشی از هوش مصنوعی ذکر شده است. همانطور که هوش مصنوعی پیچیده‌تر و گسترده‌تر می‌شود، صداها هشدار نسبت به خطرات بالقوه هوش مصنوعی بلندتر می‌شوند. **جفری هیتون، معروف به پدرخوانده هوش مصنوعی برای کار اساسی خود در زمینه شبکه عصبی و یادگیری ماشینی، گفت: "این چیزها می‌توانند هوشمندتر از ما باشند و می‌توانند تصمیم بگیرند که کنترل را در دست بگیرند، و اکنون باید نگران این باشیم که چگونه از این اتفاق جلوگیری کنیم". او شرکت گوگل را ترک کرد تا بتواند درباره خطرات هوش مصنوعی صحبت کند. او حتی گفته که از کار خودش پشیمان است**

این دانشمند مشهور کامپیوتر در نگرانی‌های خود تنها نیست. ایلان ماسک، بنیانگذار تسلا و اسپیس ایکس، به همراه بیش از ۱۰۰۰ رهبر فناوری دیگر، در نامه‌ای سرگشاده در سال ۲۰۲۳ خواستار توقف آزمایش‌های بزرگ هوش مصنوعی شدند و اشاره کردند که این فناوری می‌تواند "خطرات عمیقی" برای جامعه و بشریت ایجاد کند.

خواه این اتوماسیون فزاینده مشاغل خاص باشد، الگوریتم‌های مغرضانه جنسیتی و نژادی یا سلاح‌های مستقلی که بدون نظارت انسان عمل می‌کنند، نگرانی در تعدادی از جبهه‌ها وجود دارد. و ما هنوز در مراحل اولیه توانایی هوش مصنوعی هستیم.

اما هوش مصنوعی کامل نیست. دلایل این نقاط ضعف چه می‌توانند باشند؟:

۱- هوش مصنوعی هم‌مدل نیست

هم‌مدلی توانایی درک و به اشتراک گذاشتن احساسات دیگران است. این قابلیت است که انسان‌ها دارند و روبات‌ها ندارند.

۲- هوش مصنوعی می‌تواند افراد و فناوری را دستکاری کند

دستکاری وسایل نقلیه خودران یک نگرانی عمده عمومی است. این به این دلیل است که فناوری‌های مجهز به هوش مصنوعی کامل نیستند و می‌توان آن‌ها را دستکاری کرد.

دو نوع دستکاری ممکن است اتفاق بیفتد: اولین نوع دستکاری زمانی است که فناوری مجهز به هوش مصنوعی توسط انسان دستکاری شده و منجر به آسیب ناخواسته به انسان شود. نوع دوم دستکاری زمانی اتفاق می‌افتد که فناوری مجهز به هوش مصنوعی، سایر فناوری‌ها یا عوامل مجهز به هوش مصنوعی را دستکاری می‌کند و در نتیجه آسیب ناخواسته به افراد وارد می‌شود.

۳- فناوری‌های مجهز به هوش مصنوعی نمی‌توانند زمینه را درک کنند یا در موقعیت‌های غیرعادی قضاوت کنند.

الگوریتم‌هایی که از فناوری‌های مبتنی بر هوش مصنوعی برای تصمیم‌گیری استفاده می‌کنند، اغلب مغرضانه هستند. این به این دلیل است که داده‌هایی که آنها از آنها یاد می‌گیرند ممکن است به هر طریقی مغرضانه باشد. **به عنوان مثال، یک الگوریتم می‌تواند بر روی مجموعه داده‌ای از رزومه آموزش داده شود و بعداً متوجه شود که نسبت به متقاضیان مرد با حقوق بالاتر تعصب یا حساسیت دارد.** نوع دیگری از سوگیری الگوریتمی زمانی رخ می‌دهد که داده‌ها به روشی جانبدارانه جمع‌آوری شوند. به عنوان مثال، وقتی یک الگوریتم چیزی در مورد موضوعات خاص می‌آموزد، وقتی فقط مقالات خبری از یک طرف طیف سیاسی را می‌خواند، او تمایل به این سمت از طیف سیاسی خواهد داشت.

۴- عدم شفافیت و پاسخگویی در تصمیم‌گیری از طریق هوش مصنوعی وجود دارد

فقدان شفافیت و پاسخگویی در تصمیم‌گیری‌های هوش مصنوعی یک مشکل اساسی است. این به این دلیل است که این سیستم‌ها اغلب غیرشفاف هستند و درک نحوه تصمیم‌گیری و منطق آنها را دشوار می‌کند. این می‌تواند منجر به سوگیری، بی‌انصافی یا سایر پیامدهای ناخواسته شود. درک مدل‌های هوش مصنوعی و یادگیری عمیق، حتی برای کسانی که مستقیماً با این فناوری کار می‌کنند، دشوار است. این منجر به عدم شفافیت در مورد چگونگی و چرایی نتیجه‌گیری هوش مصنوعی می‌شود و توضیحی در مورد اینکه الگوریتم‌های هوش مصنوعی از چه داده‌هایی استفاده می‌کنند یا اینکه چرا ممکن است تصمیم‌های مغرضانه یا نالایم بگیرند، ایجاد نمی‌کند. این نگرانی‌ها منجر به استفاده از هوش مصنوعی قابل توضیح شده است، اما هنوز راه زیادی تا تبدیل شدن سیستم‌های هوش مصنوعی شفاف به روش معمول وجود دارد.

هوش مصنوعی قابل توضیح:

هوش مصنوعی قابل توضیح مجموعه‌ای از تکنیک‌ها، اصول و فرآیندهایی است که برای کمک به سازندگان و کاربران مدل‌های هوش مصنوعی استفاده می‌شود تا بفهمند این مدل‌ها چگونه تصمیم می‌گیرند. این اطلاعات می‌تواند برای بهبود دقت مدل یا شناسایی و رسیدگی به رفتارهای ناخواسته مانند تصمیم‌گیری مغرضانه استفاده شود. هوش مصنوعی قابل توضیح را می‌توان برای توصیف یک مدل هوش مصنوعی، تأثیر مورد انتظار آن و هرگونه سوگیری احتمالی، و همچنین ارزیابی دقت و انصاف آن استفاده کرد. با پیشرفت بیشتر هوش مصنوعی، بسیاری هوش مصنوعی قابل توضیح را برای آینده صنعت ضروری می‌دانند.

یادگیری عمیق:

آیا تا به حال فکر کرده‌اید که چگونه گوگل می‌تواند کل پاراگراف‌ها را از یک زبان به زبان دیگر در چند میلی‌ثانیه ترجمه کند. چگونه نتفلیکس و یوتیوب می‌توانند توصیه‌های خوبی ارائه دهند. حتی ماشین‌های خودران چگونه ممکن است؟ همه این نوآوری‌ها محصول یادگیری عمیق و شبکه‌های عصبی مصنوعی هستند.

شبکه عصبی چیست؟

شبکه عصبی مجموعه‌ای از الگوریتم‌ها است که تلاش می‌کند تا روابط زیربنایی را در مجموعه‌ای از داده‌ها از طریق فرآیندی که نحوه عملکرد مغز انسان را تقلید می‌کند، تشخیص دهد. از این نظر، شبکه‌های عصبی به سیستم‌هایی از نورون‌ها، اعم از ارگانیک یا مصنوعی اشاره دارند.

شبکه های عصبی می توانند با یک ورودی در حال تغییر سازگار شوند، بنابراین شبکه بهترین نتیجه ممکن را بدون نیاز به طراحی مجدد معیارهای خروجی ایجاد می کند.

۵- مسائل اخلاقی در ساختن یک ابرهوش

مسئله مشکلات اخلاقی با هوش مصنوعی مسئله بسیار مهمی است. امکان تولید سلاح های خودمختار نمونه بسیار خوبی از این واقعیت است. اگر هیچ قانون یا مقرراتی برای جلوگیری از توسعه این سلاح ها وجود نداشته باشد چه؟ این امکان وجود دارد که ارتش بدون نگرانی و توجه به آسیب های جانبی از آنها برای کشتن دشمن استفاده کند.

۶- هوش مصنوعی تأثیر منفی بر مشاغل و رشد اقتصادی مردم دارد

استفاده از هوش مصنوعی در محیط کار رشد تصاعدی را تجربه کرده است. تخمین زده می شود که تا سال ۲۰۲۵ هوش مصنوعی یک سوم مشاغل را به خود اختصاص دهد. جای تعجب نیست که مردم از اینکه این به چه معناست برای آینده خود و توانایی آنها برای امرار معاش می ترسند. این ترس موجه است زیرا بسیاری از صنایع در حال حاضر شاهد تأثیر اتوماسیون بر مشاغل و رشد اقتصادی هستند، کشاورزی و تولید دو بخش هستند که به شدت توسط فناوری های مبتنی بر هوش مصنوعی مختل شده اند.

۷- هوش مصنوعی خطرات امنیتی دارد

هوش مصنوعی یک فناوری نوظهور است که توانسته مشکلاتی را که قبلاً غیرقابل حل بودند حل کند. به توسعه محصولات، خدمات و فرآیندهای جدید کمک کرده است. همانطور که هوش مصنوعی پیشرفته تر می شود، می تواند راه حل هایی برای بسیاری از مشکلات در آینده ارائه دهد. با این حال، هوش مصنوعی بدون خطر نیست الگوریتم های پشتیبانی شده از هوش مصنوعی در برابر حملات هکرها آسیب پذیر هستند. این بدان معنی است که هکرها می توانند پروفایل های جعلی رسانه های اجتماعی یا برنامه های چت را برای سرقت اطلاعات از کاربران ناآگاه ایجاد کنند. اگر یک هکر موفق به سرقت اطلاعات شخصی شما شود، می تواند از آن در راه هایی استفاده کند که به شما آسیب برساند، مانند سرقت هویت یا کلاهبرداری از کارت اعتباری.

۷-۱ نگرانی های مربوط به حریم خصوصی و امنیتی:

برنامه های کاربردی هوش مصنوعی اغلب نیاز به دسترسی به مقادیر زیادی از داده ها دارند، که می تواند منجر به مسائل مربوط به حریم خصوصی یا سوء استفاده از داده ها شود و حریم خصوصی کاربر را به خطر بیندازد. خطر افشای یا هک شدن اطلاعات حساس وجود دارد - مانند داده های بهداشتی مورد استفاده در زمینه پزشکی یا ابزارهای پوشیدنی، یا داده های شخصی که توسط چت بات ها و موتورهای جستجوی هوش مصنوعی قابل دسترسی هستند.

اگر با یک چت ربات هوش مصنوعی بازی کرده باشید یا یک فیلتر چهره هوش مصنوعی را به صورت آنلاین امتحان کرده باشید، داده های شما جمع آوری می شود - اما این داده ها به کجا می رود و چگونه از آن استفاده می شود؟ سیستم های هوش مصنوعی اغلب داده های شخصی را برای سفارشی سازی تجربیات کاربر یا کمک به آموزش مدل های هوش مصنوعی که استفاده می کنید جمع آوری می کنند (مخصوصاً اگر ابزار هوش مصنوعی رایگان باشد). حتی ممکن است زمانی که داده ها به یک سیستم هوش مصنوعی داده می شوند از سایر کاربران امن تلقی نشوند، زیرا یک حادثه باگ در "چت جی پی تی" در سال ۲۰۲۳ رخ داد که به برخی از کاربران اجازه داد که عنوان تاریخچه چت کاربران فعال دیگر را ببینند. در حالی که قوانینی برای محافظت از اطلاعات شخصی در برخی موارد در ایالات متحده وجود دارد، هیچ قانون فدرال صریحی وجود ندارد که از شهروندان در برابر آسیب های حریم خصوصی داده ها که توسط هوش مصنوعی متحمل می شوند محافظت کند.

۸- فقدان دستورالعمل های اخلاقی برای اطمینان از ایمن بودن استفاده از این سیستم ها وجود دارد

فقدان دستورالعمل های اخلاقی برای اطمینان از اینکه این سیستم ها می توانند به طور ایمن استفاده شوند، یک مشکل بزرگ است. برخی از محققان استدلال می کنند که امکان ندارد نویسندگان هوش مصنوعی کاملاً عاری از سوگیری

باشند، که یک مشکل جدی است زیرا نویسندگان هوش مصنوعی اغلب در زمینه های حساس مانند مراقبت های بهداشتی، پیگیری جرایم و آموزش استفاده می شوند.

۹- بحران های مالی ناشی از الگوریتم های هوش مصنوعی

صنعت مالی پذیرای مشارکت فناوری هوش مصنوعی در فرآیندهای مالی و معاملات روزمره شده است. در نتیجه، معاملات الگوریتمی می توانند مسئول بحران مالی بعدی ما در بازارها باشد.

در حالی که الگوریتم های هوش مصنوعی تحت تأثیر قضاوت یا احساسات انسان قرار نمی گیرند، اما زمینه ها، به هم پیوستگی بازارها و عواملی مانند اعتماد و ترس انسان را نیز در نظر نمی گیرند. سپس این الگوریتم ها هزاران معامله را با سرعتی بسیار زیاد انجام می دهند و هدفشان این است که چند ثانیه بعد به سودهای کوچک بفروشند. فروختن هزاران معامله می تواند سرمایه گذاران را از انجام همین کار بترساند و منجر به سقوط ناگهانی و نوسان شدید بازار شود.

مواردی مانند سقوط فلش در سال ۲۰۱۰ و فلش نایت کاپیتال به عنوان یادآوری از نوسانات بازار ارز سخن می گوید. این بدان معنا نیست که هوش مصنوعی چیزی برای ارائه به دنیای مالی ندارد. در واقع، الگوریتم های هوش مصنوعی می توانند به سرمایه گذاران کمک کنند تا تصمیمات هوشمندانه تر و آگاهانه تری در بازار بگیرند. اما سازمان های مالی باید مطمئن شوند که الگوریتم های هوش مصنوعی و نحوه تصمیم گیری آن الگوریتم ها را درک می کنند. شرکت ها باید قبل از معرفی این فناوری برای جلوگیری از ایجاد ترس در بین سرمایه گذاران و ایجاد هرج و مرج مالی، در نظر بگیرند که آیا هوش مصنوعی اعتماد آنها را بالا می برد یا کاهش می دهد.

۱۰- از دست دادن نفوذ انسانی

اتکای بیش از حد به فناوری هوش مصنوعی می تواند منجر به از بین رفتن نفوذ انسانی - و فقدان عملکرد انسانی - در برخی از بخش های جامعه شود. به عنوان مثال، استفاده از هوش مصنوعی در مراقبت های بهداشتی می تواند منجر به کاهش همدلی و استدلال انسان شود. و استفاده از هوش مصنوعی مولد برای تلاش های خلاقانه می تواند خلاقیت و بیان احساسی انسان را کاهش دهد. تعامل بیش از حد با سیستم های هوش مصنوعی حتی می تواند باعث کاهش مهارت های ارتباطی و اجتماعی با هم تیان شود. بنابراین در حالی که هوش مصنوعی می تواند برای خودکار کردن کارهای روزانه بسیار مفید باشد، برخی این سوال را مطرح می کنند که آیا ممکن است هوش، توانایی ها و نیاز کلی انسان به جامعه را کاهش دهد یا خیر.

۱۱- سلاح های خودمختار که توسط هوش مصنوعی طراحی شده اند

همانطور که اغلب اتفاق می افتد، پیشرفت های فناوری برای اهداف جنگی به کار گرفته شده اند. وقتی صحبت از هوش مصنوعی به میان می آید، برخی تمایل دارند تا قبل از اینکه خیلی دیر شود کاری در مورد آن انجام دهند: در نامه ای سرگشاده در سال ۲۰۱۶، بیش از ۳۰۰۰۰ نفر، از جمله محققان هوش مصنوعی و رباتیک، از سرمایه گذاری در سلاح های خودران با سوخت هوش مصنوعی عقب نشینی کردند. آنها نوشتند: "سوال کلیدی برای بشریت امروز این است که آیا باید یک مسابقه تسلیحاتی جهانی هوش مصنوعی را آغاز کرد یا از شروع آن جلوگیری کرد. اگر هر قدرت نظامی بزرگی با توسعه تسلیحات هوش مصنوعی پیشروی کند، یک مسابقه تسلیحاتی جهانی عملاً اجتناب ناپذیر است و نقطه پایانی این خط سیر فناوری آشکار است: سلاح های خودمختار به کلاشینکف های فردا تبدیل خواهند شد." این پیش بینی در قالب سیستم های تسلیحاتی خودمختار مرگبار محقق شد که با رعایت مقررات کمی اهداف را به تنهایی شناسایی و نابود می کنند. به دلیل گسترش سلاح های قوی و پیچیده، برخی از قدرتمندترین کشورهای جهان تسلیم نگرانی ها شده اند و به جنگ سرد فناوری کمک کرده اند.

بسیاری از این سلاح‌های جدید خطرات عمده‌ای را برای غیرنظامیان در زمین ایجاد می‌کنند، اما این خطر زمانی تشدید می‌شود که سلاح‌های خودمختار به دست افراد اشتباه بیفتند. هکرها بر انواع مختلفی از حملات سایبری تسلط یافته‌اند، بنابراین تصور اینکه یک بازیگر مخرب به سلاح‌های خود مختار نفوذ کند و آخر زمان را تحریک کند، سخت نیست.

۱۲- هوش مصنوعی خودآگاه غیرقابل کنترل:

همچنین این نگرانی وجود دارد که هوش مصنوعی آنقدر سریع در هوش پیشرفت کند که هوشیار شود و فراتر از کنترل انسان عمل کند - احتمالاً به روشی مخرب. گزارش‌های ادعایی از این هوشیاری قبلاً رخ داده است، با یکی از گزارش‌های معروف از سوی یکی از مهندسان سابق گوگل که در یک چت بات هوش مصنوعی بیان شده، این چنان با احساس بود انگار که یک شخص با او صحبت می‌کرد. از آنجایی که نقاط عطف بزرگ بعدی هوش مصنوعی شامل ساخت سیستم‌هایی با هوش عمومی مصنوعی و در نهایت هوش مصنوعی است، بحران‌ها برای متوقف کردن کامل این تحولات همچنان در حال افزایش است.

ککک

خطرات دیگر هوش مصنوعی:

- عدم شفافیت و کنترل:

بسیاری از مدل‌های هوش مصنوعی آنقدر پیچیده هستند که درک یا کنترل نحوه عملکرد آنها تقریباً غیرممکن است. این می‌تواند رفتار آنها را غیرقابل پیش‌بینی و تصمیمات آنها را دشوار کند.

- تعصبات و تبعیض ناشی از هوش مصنوعی:

بسته به اینکه یک هوش مصنوعی با چه مقدار و چه داده‌هایی یاد می‌گیرد، خوب یا کمتر خوب کار می‌کند و همچنین می‌تواند اثرات تبعیض‌آمیز داشته باشد. به عنوان مثال، اگر با داده‌هایی که شامل تفاوت در برخورد رفتار یا رفتار تبعیض‌آمیز است، آموزش داده شود، نتایج و تصمیمات آن منعکس‌کننده همان کلیشه‌ها و تبعیض خواهد بود.

- امتناع از پذیرش سوگیری:

اگر شرکت‌ها از پذیرش سوگیری‌های ذاتی موجود در الگوریتم‌های هوش مصنوعی امتناع کنند، ممکن است ابتکارات خود را از طریق استخدام مبتنی بر هوش مصنوعی به خطر بیندازند. این ایده که هوش مصنوعی می‌تواند ویژگی‌های یک نامزد را از طریق تجزیه و تحلیل‌های چهره و صدا اندازه‌گیری کند، همچنان با تعصبات نژادی آلوده است و همان شیوه‌های استخدام تبعیض‌آمیز را بازتولید می‌کند که مشاغل ادعا می‌کنند حذف می‌کنند.

- گسترش نابرابری اجتماعی-اقتصادی:

گسترش نابرابری اجتماعی-اقتصادی ناشی از دست دادن شغل به دلیل استفاده از هوش مصنوعی یکی دیگر از دلایل نگرانی است که سوگیری‌های طبقاتی را آشکار می‌کند. کارگرانی که کارهای دستی و تکراری بیشتری انجام می‌دهند، به دلیل اتوماسیون، دستمزدشان تا ۷۰ درصد کاهش پیدا کرده است، در حالی که کارکنان اداری در مراحل اولیه هوش مصنوعی تا حد زیادی دست نخورده باقی می‌مانند. با این حال، افزایش استفاده از هوش مصنوعی مولد در حال حاضر مشاغل اداری را تحت تاثیر قرار داده است و طیف وسیعی از نقش‌ها را به وجود می‌آورد که ممکن است نسبت به سایرین در برابر دستمزد یا از دست دادن شغل آسیب‌پذیرتر باشند.

ادعاهای فراگیر مبنی بر اینکه هوش مصنوعی به نوعی بر مرزهای اجتماعی غلبه کرده یا مشاغل بیشتری ایجاد کرده است، نمی تواند تصویر کاملی از اثرات آن را ارائه دهد. در نظر گرفتن تفاوت ها بر اساس نژاد، طبقه و سایر دسته ها بسیار مهم است. در غیر این صورت، تشخیص اینکه هوش مصنوعی و اتوماسیون چگونه به ضرر افراد و گروه‌های خاص به نفع دیگران است، دشوارتر می‌شود.

اشکال مختلف سوگیری هوش مصنوعی نیز مضر هستند. اولگا روساکوفسکی، استاد علوم کامپیوتر پرینستون، در گفتگو با نیویورک تایمز گفت که تعصب هوش مصنوعی فراتر از جنسیت و نژاد است. علاوه بر داده ها و سوگیری الگوریتمی (که دومی می تواند اولی را تقویت کند)، هوش مصنوعی توسط انسان ها توسعه یافته است - و انسان ها ذاتاً سوگیری دارند. روساکوفسکی گفت: محققان عمدتاً مردانی هستند که از جمعیت‌های نژادی خاص می‌آیند، در مناطق اجتماعی-اقتصادی بالا بزرگ شده‌اند، در درجه اول افراد بدون معلولیت. ما جمعیت نسبتاً همگنی هستیم، بنابراین فکر کردن به طور گسترده در مورد مسائل جهانی یک چالش است. تجارب محدود سازندگان هوش مصنوعی ممکن است توضیح دهد که چرا هوش مصنوعی تشخیص گفتار اغلب لهجه‌ها و لهجه‌های خاص را درک نمی‌کند، یا چرا شرکت‌ها عواقب جعل چهره‌های بدنام تاریخ بشریت است در یک ربات چت در نظر نمی‌گیرند. توسعه‌دهندگان و کسب‌وکارها باید دقت بیشتری به خرج دهند تا از بازآفرینی سوگیری‌ها و پیش‌داوری‌های قدرتمندی که جمعیت اقلیت‌ها را در معرض خطر قرار می‌دهد، اجتناب کنند.

- مشکلات کیفیت و حساسیت به خطاها:

سیستم های هوش مصنوعی بدون خطا نیستند و می توانند نتایج نادرست یا گمراه کننده ای ارائه دهند، به خصوص اگر با داده های آموزشی ناکافی یا تحریف شده کار کنند یا برنامه ریزی شده اند تا در هر موردی پاسخی بدهند - حتی اگر داده ها این اجازه را ندهند و بنابراین باید پاسخ ها را ابداع کرد. بنابراین اطلاعات ربات‌های چت یا موتورهای جستجوی هوش مصنوعی باید همیشه با چندین منبع مقایسه شوند.

- واقعیت تحریف شده و حباب‌های فیلتر:

در رسانه‌های اجتماعی، الگوریتم‌های مبتنی بر هوش مصنوعی تضمین می‌کنند که فقط محتوای ظاهراً مرتبط به ما نشان داده می‌شود. این می‌تواند منجر به ایجاد حباب‌های فیلتر شود که درک یک طرفه را ترویج می‌کند. ابزارهای هوش مصنوعی مانند فیلترهای زیبایی که تصاویر غیرواقعی بدن را تقویت می‌کنند و در نتیجه بر زنان و جوانان به‌ویژه فشار وارد می‌کنند، می‌توانند به تحریف واقعیت کمک کنند.

- سوء استفاده و دستکاری:

برنامه های کاربردی هوش مصنوعی همچنین می توانند برای اهداف سوء استفاده یا دستکاری استفاده شوند، به عنوان مثال. ایجاد دیپ فیک، انتشار اطلاعات نادرست یا نظارت و کنترل کاربران خود. اگر از هوش مصنوعی برای تبلیغات سیاسی یا دستکاری اجتماعی استفاده شود، می‌تواند منجر به مشکلات جدی فردی و اجتماعی شود.

دستکاری اجتماعی نیز به عنوان یک خطر هوش مصنوعی به حساب می‌آید. این ترس به واقعیت تبدیل شده است، زیرا سیاستمداران برای ترویج دیدگاه‌های خود به پلتفرم‌ها متکی هستند، به عنوان مثال فردیناند مارکوس جوان در جریان انتخابات فلیپین در سال ۲۰۲۲ **ارتش ترول** را با کمک "تیک توک" بکار گرفت.

حال این ارتش از کجا آمده؟

نام رایج برای یک سازمان مخفی در روسیه که اینترنت را از طرف دولت دستکاری می‌کند. با کمک عروسک‌های جورابی - هویت‌های ساختگی - روحیه عمومی در انجمن‌های آنلاین و بخش نظرات سایت‌های خبری تحت تأثیر منافع دولت روسیه قرار می‌گیرد.

نحوه تشخیص دیپ فیک:

دیپ فیک‌ها عکس‌ها، فیلم‌ها یا فایل‌های صوتی دستکاری شده‌ای هستند که با استفاده از هوش مصنوعی ایجاد شده‌اند. آنها افرادی را نشان می‌دهند که کارهایی را انجام می‌دهند یا می‌گویند که هرگز انجام نداده‌اند یا نگفته‌اند. این اصطلاح از دو کلمه "یادگیری عمیق" (روشی که هوش مصنوعی از طریق آن یاد می‌گیرد) و "جعلی" ساخته شده است (دیپ فیک) دیپ فیک اغلب برای تبلیغات سیاسی و همچنین برای مقاصد پورنوگرافی استفاده می‌شود. آنها را می‌توان به راحتی با استفاده از برنامه‌ها یا فیلترهای داخلی ایجاد کرد. با این حال، استفاده از آنها بدون رضایت شخص مورد نظر ممکن است حق چاپ را نقض کند.

برخی ویژگی‌های معمولی وجود دارد که می‌تواند نشان‌دهنده دیپ‌فیک باشد: این ویژگی‌ها عبارتند از چهره غیرطبیعی، انتقال‌های تار، عدم نیز می‌تواند به شما در تشخیص دیپ فیک کمک چشمک زدن، کیفیت ناسازگار و پس‌زمینه غیرمنطقی. ابزارهایی هم هستند که می‌توانند به کنند. با این حال، از آنجایی که فناوری‌های موجود در پشت دیپ فیک‌ها به پیشرفت خود ادامه می‌دهند، همیشه باید زمینه چنین محتوای آنالین را بررسی کنید.

گزارشات نادرست گزارش‌ها، مقالات، پست‌ها یا ویدیوهای حاوی اطلاعات نادرست هستند که عمدتاً در اینترنت منتشر می‌شوند. هدف فریب دادن، ترساندن، بیقراری هر چه بیشتر مردم یا ایجاد احساسات علیه یک علت یا گروه جمعیتی خاص است. این روزها اخبار جعلی به سرعت در شبکه‌های اجتماعی پخش می‌شود.

نمونه‌هایی از اخبار جعلی در اینترنت:

- جعل و گزارش‌های طنز: این شامل گزارش‌های طنز و همچنین شایعات صرف است که اغلب تشخیص آنها از اخبار واقعی دشوار است. حتی اگر قصد طنز یا شوخ طبعی داشته باشند، می‌توانند مردم را فریب دهند و عمدتاً گمراه کنند.
- رقابت‌ها و پیشنهادات تقلبی: عناوین نادرستی که وعده جایزه خاص یا دسترسی به پیشنهادات انحصاری را می‌دهند، کاربران را فریب می‌دهند تا اطلاعات شخصی خود را افشا کنند یا لینک‌های مضر کلیک کنند. سوء استفاده از داده‌ها یا آسیب مالی می‌تواند پیامدهای چنین تلاش‌های کلاهبرداری باشد.
- نظرات و گواهی‌های کارشناسی جعلی: ادعاهای نادرستی که ظاهراً از سوی کارشناسان یا شخصیت‌های شناخته شده می‌آید، از اعتماد کاربران به منظور تبلیغ برخی محصولات، خدمات یا دیدگاه‌های سیاسی سوء استفاده می‌کند.
- گزارش‌ها و رسوایی‌های جنجالی نادرست: تیترهای هیجان‌انگیزی که رویدادهای هیجان‌انگیز یا جنجالی را اغراق یا تحریف می‌کنند و توجه خوانندگان را به خود جلب می‌کنند. هدف ممکن است برانگیختن ترس یا تضعیف اعتماد به افراد یا مؤسسات خاص باشد.
- گزارش‌های دستکاری و پیام‌های نفرت‌انگیز: این پیام‌ها عمدتاً در شبکه‌های اجتماعی پخش می‌شوند و با هدف ایجاد برای یا علیه یک هدف سیاسی، گروه‌های جمعیتی خاص یا افراد هستند. برخی از این اخبار جعلی بر اساس رویدادهای واقعی هستند که تحریف یا اشتباه جلوه داده شده‌اند، در حالی که برخی دیگر کاملاً ساختگی هستند.

چگونه بدانیم که اخبار جعلی هستند؟

- علامت‌ها و نشانه‌هایی وجود دارند که اشاره به اخبار جعلی می‌کنند، مانند:
 - زبان مبهم و هیجان‌انگیز (مثلاً «خدای من! عالی! باور نکردنی!»، «باور نمی‌کنید بعد چه اتفاقی افتاد!»)
 - اصطلاحات احساسی (مانند "وحشتناک"، "دلخراش"، "باور نکردنی")
 - عناوین غیر طبیعی: کلاهبرداری‌ها اغلب دارای عناوین پوچ یا اغراق‌آمیز هستند که برای جلب توجه طراحی شده‌اند.
 - تصاویر هوش مصنوعی شدید و تکان‌دهنده
 - فقدان جزئیات یا منابع: حقه‌ها اغلب حاوی جزئیات یا منابع کمی هستند یا هیچ منبعی برای حمایت از ادعاهای خود ندارند.
 - اعداد و آمار مشکوک بدون مدرک
 - حساب‌های مشکوک: گزارش‌های نادرست اغلب توسط حساب‌ها یا حساب‌های جعلی بدون تأیید پخش می‌شوند (مثلاً تیک آبی در اینستاگرام).
 - برانگیختن ترس (مثلاً «این یک تهدید برای همه ماست!»، «می‌بینید»)
 - تعمیم‌ها (مثلاً «همه... هستند...»)
- تئوری‌های توطئه (به عنوان مثال در مورد مغز متفکران مخفی، ذینفعان، قدرت‌های ناشناخته، و غیره)

بدین ترتیب در برابر اخبار جعلی اقدام کنید:

گزارش‌های نادرست را نادیده بگیرید و آنها را بیشتر پخش نکنید. پیام‌های مشکوک را حذف یا نادیده بگیرید، عجولانه روی لینک‌ها کلیک نکنید یا پیوست‌های ناشناخته را باز نکنید. مهمتر از همه، گزارش‌های نادرست را فوراً رد نکنید، لایک نکنید یا به اشتراک نگذارید. افرادی که اخبار نادرست را منتشر می‌کنند یا پست‌های خود را مخفی می‌کنند، مسدود کنید

- سخنرانی مخالف ارائه دهید و به طور عینی بحث کنید. با بررسی محتوای آنلاین و اصلاح آن با حقایق، گزارش های نادرست را رد کنید. ضد گفتار می تواند ابزار قدرتمندی برای مقابله با ادعاهای نادرست و نفرت آنلاین باشد. اما همیشه عینی و محترمانه بحث کنید و اجازه ندهید که تحریک شوید.

- تقلب ها را گزارش کنید. می توانید گزارش های نادرست و پست های نفرت انگیز را مستقیماً در شبکه های اجتماعی و دفاتر گزارش رسمی گزارش دهید. برای محدود کردن گسترش آنها باید حتماً این کار را انجام دهید.

نکاتی برای معلمان در مورد نحوه استفاده از هوش مصنوعی در کلاس درس:

نادیده نگیرید. دانش آموزان به هر صورت از ابزار " چت جی بی تی " استفاده می کنند با اینکه شما در مقابلش شک و شبه دارید. بنابراین خود را در مقابل فرصت و ریسک های برنامه های کاربردی هوش مصنوعی قرار داده و نشان دهید که نقد منابع چه جایگاهی دارد. - کلا ممنوع نکنید. ممنوعیت فرصت تعامل فعال با موضوع و همچنین اشاره به نقاط ضعف هوش مصنوعی را از بین می برد. - مشترکاً استفاده کنید. با یکدیگر امکانات و محدودیت های برنامه های هوش مصنوعی و قوانین طراحی برای زندگی روزمره را در کلاس و مدرسه کشف کنید.

نتیجه استفاده از هوش مصنوعی در مدارس تا کنون:

استفاده از هوش مصنوعی در مدارس هنوز با تردید مواجه است. این واقعیت که دانش آموزان به طور فزاینده ای از برنامه های کاربردی هوش مصنوعی استفاده می کنند، ارزیابی عملکرد را برای معلمان دشوارتر می کند، در حالی که دانش آموزان اغلب تشخیص پاسخ های نادرست از چت ربات های هوش مصنوعی و ارزیابی صحیح اطلاعات را دشوار می کنند. این امر مشارکت فعالانه با هوش مصنوعی در کلاس و تمرکز بیشتر بر ارتقای سواد اطلاعاتی و تمرین انتقاد منبع را بیش از پیش مهم می کند.

حجج

