



## ۵ رویداد مهم هوش مصنوعی در سال ۲۰۲۴ کدامند؟

انتظار می‌رود که هوش مصنوعی در سال ۲۰۲۴ رویدادهای مهمی را پشت سر بگذارد. در این گزارش به بررسی پنج رویداد مهم می‌پردازیم.

انتظار می‌رود که هوش مصنوعی در سال ۲۰۲۴ رویدادهای مهمی را پشت سر بگذارد. در این گزارش به بررسی پنج رویداد مهم می‌پردازیم.

به گزارش ایسنا و به نقل از فوربس، سال ۲۰۲۳ یک نقطه عطف در تکامل فناوری بود و هوش مصنوعی مولد به جریان اصلی آن تبدیل شد. با آغاز سال ۲۰۲۴ انتظار می‌رود که چشم انداز هوش مصنوعی مولد به سرعت در حال تکامل یافتن باشد و مجموعه ای از رویدادها را پشت سر بگذارد که تغییر فناوری و کاربردهای آن را نوید می‌دهند.

این رویدادها از پیشرفت در مدل‌های هوش مصنوعی چندوجهی گرفته تا ظهور مدل‌های زبانی کوچک، نه تنها چشم انداز فناوری را شکل می‌دهند، بلکه تعاملات، خلاقیت و درک پتانسیل هوش مصنوعی را دوباره تعریف می‌کنند. در این گزارش، برترین رویدادهای هوش مصنوعی مولد در سال ۲۰۲۴ را بررسی می‌کنیم.

### ۱. ظهور مدل‌های هوش مصنوعی چندوجهی

چت بات GPT۴ شرکت «OpenAI»، «لاما ۲» (LLama ۲) شرکت متا و هوش مصنوعی شرکت «میسترال ای آی» (Mistral AI) همگی نمونه‌هایی از پیشرفت در مدل‌های زبانی بزرگ بودند. این فناوری با کمک مدل‌های هوش مصنوعی چندوجهی، فراتر از متن حرکت می‌کند و به کاربران امکان می‌دهد تا محتوا را براساس متن، صدا، تصویر و ویدیو برای ایجاد محتوای جدید ترکیب کنند و مطابقت دهند. این روش شامل ترکیب داده‌هایی مانند تصویر، متن و گفتار با الگوریتم‌های پیشرفته برای پیش‌بینی و ایجاد نتایج است.

انتظار می‌رود در سال ۲۰۲۴، هوش مصنوعی چندوجهی به طور قابل توجهی تکامل یابد و تغییراتی را در قابلیت‌های هوش مصنوعی مولد ایجاد کند. این مدل‌ها فراتر از عملکردهای تک‌حالتی سنتی پیش می‌روند و داده‌های متنوعی مانند تصویر، زبان و صدا را در خود جای می‌دهند. هوش مصنوعی در نتیجه انتقال یافتن به مدل‌های چندوجهی، شهودک‌تر و پویاتر می‌شود. چت بات «GPT۴-V» در حال حاضر میان مشترکین «ChatGPT Plus» به دلیل قابلیت‌های چندوجهی خود محبوب است. در سال ۲۰۲۴ می‌توان انتظار ظهور مدل‌های منبع باز مانند «LLava» را داشت.

### ۲. مدل‌های زبانی کوچک توانا و قوی

اگر سال ۲۰۲۳ سال مدل‌های زبان بزرگ بود، سال ۲۰۲۴ شاهد قدرت‌نمایی مدل‌های زبانی کوچک است. مدل‌های زبانی بزرگ با مجموعه بزرگی از داده‌ها مانند «کامن کراول» (Common Crawl) و «دیل» (The Pile) آموزش می‌بینند. داده‌های این مجموعه‌ها از میلیاردها وبسایت در دسترس عموم استخراج شده‌اند. اگرچه داده‌های مورد استفاده در آموزش مدل‌های زبانی بزرگ برای تولید محتوای معنادار و پیش‌بینی کلمه بعدی واقعا سودمند هستند اما ماهیت آنها از محتوای عمومی اینترنت حاصل می‌شود.

از سوی دیگر، مدل‌های زبانی کوچک با مجموعه داده‌های محدودتری آموزش داده می‌شوند. با وجود این، آنها از منابع باکیفیت مانند کتاب‌های درسی، مجلات و محتوای معتبر تشکیل شده‌اند. این مدل‌ها از نظر تعداد پارامترها و همچنین نیازهای ذخیره‌سازی و حافظه کوچک‌تر هستند و این موجب می‌شود که آنها با سخت‌افزارهای کم‌قدرت و ارزان‌تری کار کنند. مدل‌های زبانی کوچک با وجود این که کسری از اندازه مدل‌های زبانی بزرگ را دارند اما کیفیت محتوای آنها با برخی از همتایان بزرگ‌تر قابل مقایسه است.

مدل‌های «PHI-۲» و «میسترال ۷بی» (Mistral VB) شرکت مایکروسافت، دو مدل زبانی کوچک امیدوارکننده هستند که قدرت نسل بعدی برنامه‌های کاربردی هوش مصنوعی را تأمین می‌کنند.

### ۳. ظهور عوامل خودمختار

عوامل خودمختار یک راهبرد نوآورانه برای ساخت مدل‌های هوش مصنوعی مولد هستند. این عوامل، برنامه‌های نرم‌افزاری مستقلی هستند که برای دستیابی به یک هدف خاص طراحی شده‌اند. هنگام در نظر گرفتن هوش مصنوعی مولد، توانایی عوامل مستقل برای تولید محتوای عاری از مداخله انسانی بر محدودیت‌های مرتبط با مهندسی مرسوم غلبه می‌کند. الگوریتم‌های پیشرفته و روش‌های یادگیری ماشینی برای توسعه عوامل مستقل استفاده می‌شوند. این عوامل از داده‌ها برای یادگیری، سازگاری با موقعیت‌های جدید و تصمیم‌گیری با دخالت اندک انسان استفاده می‌کنند. به عنوان مثال، OpenAI مدل‌هایی را ابداع کرده است که از عوامل مستقل استفاده موثری می‌کنند و نشان دهنده پیشرفت قابل توجه در زمینه هوش مصنوعی هستند.

هوش مصنوعی چندوجهی که ترکیبی از روش‌های گوناگون هوش مصنوعی مانند پردازش زبان طبیعی، بینایی رایانه‌ای و یادگیری ماشینی است، اهمیت بسیاری در توسعه عوامل مستقل دارد. هوش مصنوعی چندوجهی می‌تواند با تحلیل داده‌های گوناگون به طور هم‌زمان و اعمال مفاهیم کنونی پیش‌بینی کند، اقداماتی را انجام دهد و تعامل مناسب‌تری داشته باشد.

چارچوب‌هایی مانند «لنگ چین» (LangChain) و «لاما ایندکس» (LlamaIndex) برخی از مدل‌های محبوبی هستند که برای ساخت عوامل براساس مدل‌های زبانی بزرگ استفاده می‌شوند. در سال ۲۰۲۴ شاهد ظهور چارچوب‌های جدیدی خواهیم بود که از هوش مصنوعی چندوجهی بهره می‌برند.

#### ۴. مدل‌های منبع باز قابل قیاس با مدل‌های انحصاری

انتظار می‌رود که در سال ۲۰۲۴، مدل‌های هوش مصنوعی باز و مولد به طور قابل توجهی تکامل پیدا کنند و برخی از پیش‌بینی‌ها نشان می‌دهند که آنها با مدل‌های انحصاری قابل مقایسه خواهند بود. از سوی دیگر، مقایسه بین مدل‌های باز و انحصاری پیچیده است و به عوامل گوناگونی از جمله موارد استفاده خاص، منابع توسعه و داده‌های مورد استفاده برای آموزش مدل‌ها بستگی دارد.

مدل‌های «لاما ۲ ۷۰ بی» (Llama ۲ ۷۰B)، «فالکون ۱۸۰ بی» (Falcon ۱۸۰B) و «میس‌ترال ۸ ایکس ۷ بی» (Mixtral 8x7B) شرکت میس‌ترال ای‌آی در سال ۲۰۲۳ بسیار محبوب شدند و عملکرد آنها با مدل‌های انحصاری مانند «GPT ۳.۵»، «کلاد ۲» (Claude ۲) و «ژوراسیک» (Jurassic) قابل مقایسه بود.

در آینده، شکاف بین مدل‌های باز و مدل‌های انحصاری کاهش می‌یابد و یک گزینه عالی را برای میزبانی مدل‌های هوش مصنوعی در محیط‌های ترکیبی یا داخلی به شرکت‌ها ارائه می‌کند.

مدل‌های بعدی متا، میس‌ترال ای‌آی و احتمالاً شرکت‌های جدید در سال ۲۰۲۴ به عنوان جایگزین‌های مناسبی برای مدل‌های انحصاری موجود به عنوان واسط برنامه نویسی کاربردی ارائه خواهند شد.

#### ۵. «ابر بومی» کلید رسیدن به هوش مصنوعی مولد

پلتفرم «کوبرنتیز» (Kubernetes) در حال حاضر یک محیط ترجیحی برای میزبانی مدل‌های هوش مصنوعی مولد است. انتظار می‌رود شرکت‌های مهمی مانند «هاگینگ فیس» (Hugging Face)، گوگل و OpenAI از زیرساخت‌های «ابر بومی» (Cloud Native) که توسط کوبرنتیز پشتیبانی می‌شوند، برای ارائه پلتفرم‌های هوش مصنوعی مولد استفاده کنند.

مدل‌هایی مانند «تکست جنریشن اینفرنس» (Text Generation Inference) شرکت هاگینگ فیس، «ری سرو» (Ray Serve) شرکت «انی اسکیل» (AnyScale) و «vLLM» در حال حاضر از اجرای زیرساخت‌های مورد نظر پشتیبانی می‌کنند. در سال ۲۰۲۴ شاهد بلوغ چارچوب‌ها، مدل‌ها و پلتفرم‌هایی خواهیم بود که در کوبرنتیز اجرا می‌شوند تا کل چرخه عمر مدل‌های پایه را مدیریت کنند. کاربران می‌توانند مدل‌های تولیدی را تحت پیش‌آموزش، تنظیم دقیق، استقرار و مقیاس بندی کارآمد قرار دهند.