

«دیپ‌مایند» چگونه با حل مساله ساختار پروتئین‌ها دنیای پزشکی را دگرگون می‌کند؟ - دیجیاتو

آکام شبرندی | چهارشنبه، ۱۲ آذر ۱۳۹۹

اتفاقی که همه چیز را تغییر می‌دهد؛ این خبری بود که دیروز به تیتراژ اول بسیاری از نشریات علمی و معتبر از جمله «نیچر» و «ساینس» بدل شده و از حل یک مساله ۵۰ ساله زیست‌شناسی توسط هوش مصنوعی حکایت داشت. برای آشنایی بیشتر با این دستاورد بزرگ که انقلابی در زمینه زیست‌شناسی مولکولی بر پا کرده در ادامه مطلب با ما همراه باشید

«[دیپ مایند](#)»، یکی از پیشگام‌ترین کمپانی‌های حوزه هوش مصنوعی است که از زیرمجموعه‌های [گوگل](#) محسوب می‌شود و اخیراً به لطف هوش مصنوعی موفقیت چشمگیری را در پیش بینی ساختار پروتئین‌ها کسب کرده است. این شرکت [دیروز اعلام کرد](#) که با سیستم «آلفا فولد» خود چالش بزرگ چینش و تاشدگی پروتئین را حل کرده که به مدت 50 سال جامعه علمی را ناامید کرده بود.

پیش‌بینی ساختار سه بعدی پروتئین برای نیم‌قرن یک چالش بزرگ بوده است. هوش مصنوعی این شرکت توانسته یکی از پیچیده‌ترین شبیه‌سازی‌های ساختار پروتئین‌ها را که قبلاً به صورت تجربی تنها در آزمایشگاه‌های عظیم انجام شده، این بار با هوش مصنوعی عملی کند. این الگوریتم در واقع شکل سه بعدی پروتئین‌ها را بوسیله توالی اسید آمینه‌های آن پیش بینی کرده و جهشی عظیم در حل یکی از بزرگترین چالش‌های زیست‌شناسی محسوب می‌شود.

نتایج تحقیقات این شرکت در آغاز کنفرانس دوسالانه «CASP» اعلام شد. در این کنفرانس 100 تیم شرکت کننده دیگر نیز حضور داشتند اما نتایج پروژه دیپ‌مایند از بقیه بهتر بود. «جان مولت»، استاد دانشگاه «مریلند»، که در سال 1994 این کنفرانس را با هدف بهبود روش‌های محاسباتی برای پیش بینی دقیق ساختارهای پروتئینی بنیان نهاد، می‌گوید: «این اتفاق مهمی است و می‌شود گفت به معنی واقعی مشکل حل شده است.»



توانایی پیش بینی دقیق ساختارهای پروتئینی از توالی اسید آمینه آنها یک پیشرفت بزرگ برای علوم زیستی و پزشکی است و باعث تسریع تلاش‌های دانشمندان در شناخت کامل سلول‌ها و حتی تولید داروهای پیشرفته می‌شود. برای نمونه می‌توان به داروهای منوکلونال آنتی بادی اشاره کرد که به غشای سلول‌های هدف اتصال پیدا می‌کنند. این داروها به صورت مستقیم توانایی شناسایی یک مولکول خاص را دارند و به همان مولکول اتصال می‌یابند و از این طریق در درمان بسیاری از سرطان‌ها و حتی بیماری‌های خود ایمنی کاربرد دارند؛ برای مثال داروی «رتاکسیمب» که در درمان بیماری MS کاربرد دارد.

این داروها با توجه به ساختار سه بعدی که دارند اختصاصا به یک گیرنده خاص در سلول هدف متصل شده و باعث ایجاد اثر درمانی می‌شوند. برای تولید اینگونه داروها حتما باید ساختار سه بعدی پروتئین گیرنده و ساختار سه بعدی دارو را بدون هیچگونه اشتباهی تعیین کنیم. همچنین درک ساختار سه بعدی پروتئین‌ها و آمینو اسیدها در ساختار این داروها و حتی پروتئین‌های هدف در این مورد لازم است و اینجاست که این عظمت این پروژه و کارایی آن در کمک به دانشمندان و پزشکان برای تشخیص و درمان بیماری‌ها برای ما نمایان می‌شود.

تیم مستقر در لندن «آلفافولد» در سال 2018 برای اولین بار در این کنفرانس شرکت کرد و امسال یعنی سال 2020 از تمام رقبا یک سر و گردن جلوتر بود و حتی به گفته دانشمندان در حال ایجاد یک انقلاب در زمینه زیست شناسی است. «آندری لوپاس»، زیست شناس تکاملی در موسسه توسعه زیست شناسی ماکس پلانک در توبینگن آلمان، که عملکرد تیم‌های مختلف را در CASP ارزیابی کرد بر این باور است که «این اتفاق بازی را تغییر می‌دهد.»

هوش مصنوعی پیش‌بینی کننده دیپمایند آلفافولد نام دارد
تیم آلفا فولد به او کمک کرده است تا ساختار پروتئینی را که به مدت یک دهه در آزمایشگاه به

دنبال آن بوده و او را گنج کرده بود، پیدا کند. از همین رو انتظار دارد این پیشرفت باعث تغییراتی اساسی در نحوه مطالعات و دانش پزشکی شود: «این اتفاق پزشکی، تحقیقات و مطالعات، مهندسی زیستی و همه چیز را تغییر خواهد داد.»

البته این همه داستان نیست، گاهی اوقات ساختار پیش بینی شده هوش مصنوعی گوگل با ساختار اصلی و واقعی که به صورت تجربی و برپایه روش‌های «استاندارد طلایی» مانند کریستالوگرافی اشعه ایکس و یا میکروسکوپ الکترونی مشخص شده، متفاوت است.

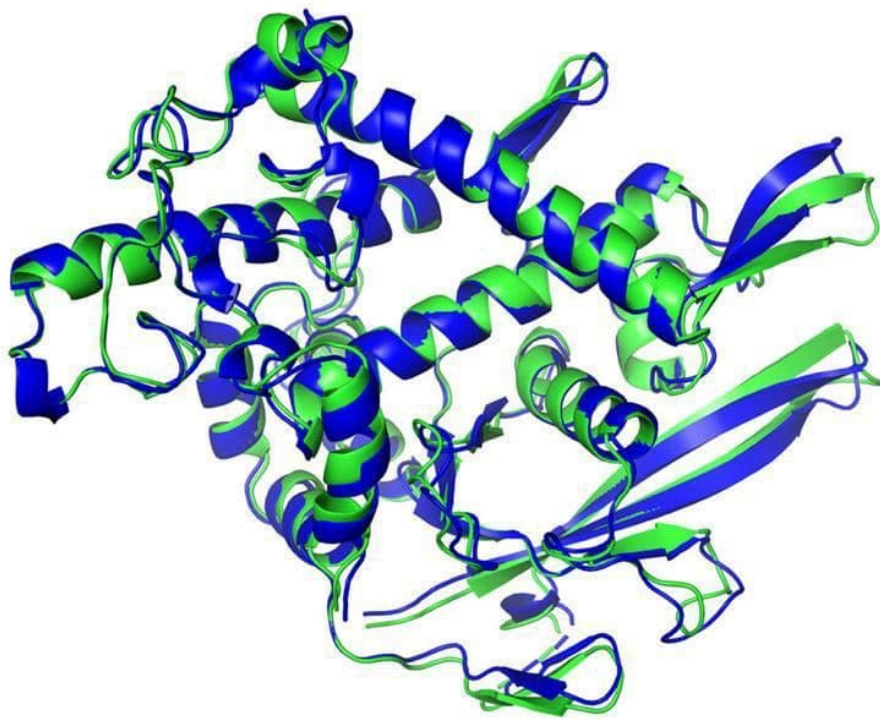
این موضوع نشان می‌دهد نیاز ما به این روش‌های گران قیمت ممکن در است در سال‌های آینده از بین نرود. اما هوش مصنوعی به ما کمک می‌کند که روی موجودات زنده با روش‌های جدید مطالعه کنیم.

مشکل ساختار

پروتئین‌ها عناصر سازنده زندگی هستند و مسئولیت بیشتر آنچه در داخل سلول‌ها اتفاق می‌افتد را بر عهده دارند. نحوه کار و عملکرد پروتئین‌ها توسط شکل سه بعدی آن‌ها تعیین می‌شود. هر پروتئین توسط یک ژن تولید می‌شود. البته مطالعات جدید نشان داده هر ژن در واقع مسئول تولید یک رشته آمینو اسیدی است.

این یکی از قوانین زیست شناسی مولکولی این است که «ساختار عملکرد است». در واقع پروتئین‌ها تمایل دارند بدون کمک، شکل تعیین شده خود را بدست آورند. این موضوع توسط قوانین فیزیک هدایت می‌شود و شکل هر پروتئین با توجه به ساختار آمینو اسیدی و توالی آن‌ها مشخص می‌شود.

برای دهه‌ها، تجربیات آزمایشگاهی اصلی‌ترین و تنها راه برای تعیین ساختارهای پروتئینی بود. اولین ساختارهای کامل پروتئین‌ها با تکنیکی کشف شد که در آن پرتوهای اشعه ایکس به پروتئین‌های بلوری شکل (با روش‌های پیچیده شیمیایی پروتئین‌ها بلوری شکل می‌شوند) شلیک می‌شدند و نور انعکاس یافته به مختصات اتمی پروتئین تبدیل می‌شد. به این روش کریستالوگرافی اشعه ایکس می‌گویند.



شمایی از یک ساختار سه بعدی پروتئینی که قسمت آبی توسط هوش مصنوعی و قسمت سبز با روش‌های تجربی تعیین شده. پیش‌بینی‌ها تقریباً باهم تطابق کامل دارند.

کریستالوگرافی اشعه ایکس سهم مهمی در کشف ساختارهای پروتئینی ایفا کرده اما طی دهه گذشته، Cryo-EM (با استفاده از میکروسکوپ الکترونی) به ابزار مورد علاقه بسیاری از آزمایشگاه‌های زیست‌شناسی ساختاری تبدیل شده است. همانطور که در بالا اشاره کردیم این ابزارها استاندارد طلایی تشخیص ساختارهای پروتئینی هستند.

این که پروتئین‌ها چگونه با اجزای ساختاری یعنی اسیدهای آمینه شکل خود را تعیین می‌کنند برای محققان، مدت‌ها جای سوال بود. تلاش‌های اولیه استفاده از رایانه برای پیش‌بینی ساختارهای پروتئینی در دهه 1980 و 1990 عملکرد ضعیفی داشتند. بسیاری از ادعاها در این زمینه زمانی که محققان آن‌ها را با روش‌های علمی بررسی می‌کردند رد می‌شد و این روش‌ها قدرت پیش‌بینی ضعیفی داشتند.

آلفافولد بین ۱۰۰ رقیب توانسته برای اولین بار به امتیاز بالای ۹۰ در پیش‌بینی برسد دلیل اصلی جان مولت هم برای راه‌اندازی CASP سخت‌گیری بیشتر در مورد این دست ادعاها بود. این سلسله کنفرانس‌ها تیم‌ها را به چالش می‌کشند تا ساختار پروتئین‌هایی را که با استفاده از روش‌های آزمایشی (استاندارد طلایی) حل شده‌اند، اما ساختارهای آنها عمومی نشده، پیش‌بینی کنند. مولت این کنفرانس را یک فضای رقابتی نمی‌بیند و می‌گوید:

«شما خواهید فهمید چه چیزهایی امیدوارکننده به نظر می‌رسد، چه چیزهایی جواب می‌دهد و باید از چه چیزهایی دور شوید.»

هرچند عملکرد دیپ‌ماینر در CASP13 که دو سال قبل برگزار شد آنچنان درخشان نبود اما بسیاری از دانشمندان این حوزه را که مدتها بود در گروه‌های کوچک دانشگاهی موفقیتی نداشتند، مبهور کرد. «جان جمپر» که رهبری گروه DeepMind را به عهده دارد می‌گوید:

«در ابتدا با هوش مصنوعی فاصله آمینو اسیدهای جفت شده سنجیده می‌شد و سپس در قسمت بعدی که از هوش مصنوعی استفاده نمی‌شد نرم افزارها با داده‌های حاصل از مرحله اول مشخص می‌کردند که شکل پروتئین باید چگونه باشد.»

تیم آلفا فولد سعی در ادامه این روش داشت اما در نهایت شکست خورد. جمپر می‌گوید، بنابراین این روند را تغییر دادیم و یک شبکه هوش مصنوعی ایجاد کردیم که شامل اطلاعات اضافی در مورد محدودیت‌های فیزیکی و هندسی انواع پروتئین‌ها بود که نحوه چین خوردگی آن‌ها را تعیین می‌کرد. آنها همچنین کار سخت‌تری را باید انجام می‌دادند: شبکه به جای پیش بینی روابط بین اسیدهای آمینه، ساختار نهایی پروتئین هدف را پیش بینی می‌کرد؛ سیستمی که به باور جامپر بسیار پیچیده تر است

دقت حیرت انگیز

کنفرانس دوسالانه CASP طی چندین ماه انجام می‌شود. پروتئین‌های هدف یا بخشی از پروتئین‌ها که به آن نام «دامین» داده‌اند (در مجموع حدود 100 نوع پروتئین) به طور منظم در طی چندین هفته به تیم‌ها داده می‌شوند و تیم‌ها چندین هفته فرصت دارند تا پیش بینی‌های خود را در رابطه با ساختار پروتئین مورد نظر ارائه دهند. سپس تیمی از دانشمندان مستقل پیش بینی‌های ارائه شده از تیم‌ها را با استفاده از معیارهای تعیین شده از طریق روش‌های استاندارد طلایی اندازه گیری می‌کنند. داوران نمی‌دانند چه کسی ساختارها را پیش بینی کرده است و داوران‌ها به صورت کور یا بلیند انجام می‌شود.

دستاورد دیپ‌ماینر در تشخیص و درمان بیماری‌ها تاثیر عمیقی خواهد داشت لوپاس می‌گوید، پیش بینی‌های آلفا فولد با نام "گروه 427" وارد مسابقه شده بود و به واسطه دقت حیرت انگیز بسیاری از پیش بینی‌های آنها بین دوران شهرت پیدا کردند. وی می‌گوید: «حدس می‌زدیم که آنها باید آلفا فولد باشند، بیشتر داوران هم چنین حدسی داشتند.»

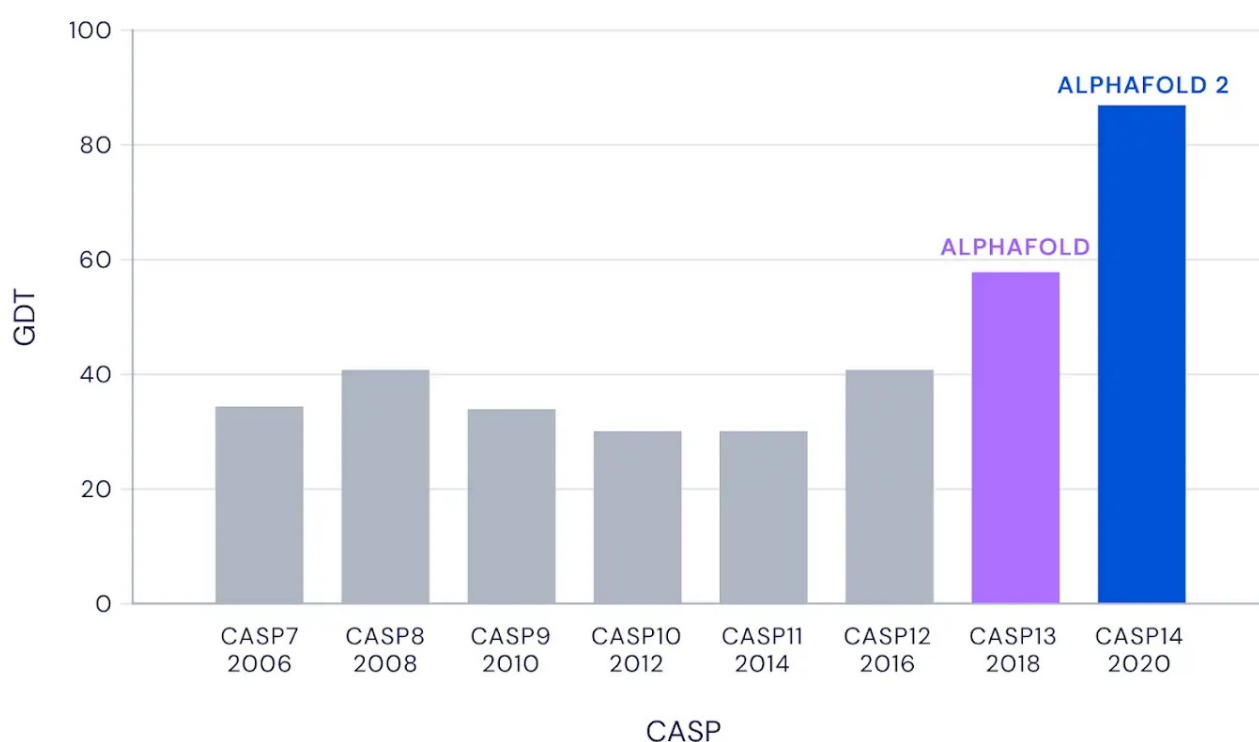
برخی از پیش بینی‌ها از بقیه بهتر بود، اما تقریباً دو سوم پیش بینی‌ها از نظر کیفیت در حد ساختارهای آزمایشی (استاندارد طلایی) بودند. دقت روش‌های تیم آلفا فولد در حدی بود که حتی مولت معتقد بود اختلاف‌های بین ساختارهای پیش بینی شده و تعیین شده چنان کم بود که گاهی اوقات مشخص نبود اختلاف از طرف خطاهای محاسباتی آلفا فولد پیش آمده و یا خطاهایی در روش‌های استاندارد رخ داده است. مولت در ادامه می‌گوید:

«بعضی از پیش بینی‌های آلفا فولد با ساختارهای آزمایشی تعیین شده توسط تکنیکی به نام

طیف سنجی تشدید مغناطیسی هسته‌ای مطابقت نداشت، اما این می‌تواند به چگونگی تبدیل داده‌های خام به مدل ساختاری پروتئینی ارتباط داشته باشد.»

این شبکه حتی تعامل گروهی پروتئین‌ها بر هم را بررسی می‌کند و عملکرد هر پروتئین را هم به صورت مجزا و هم به صورت گروهی پیش بینی می‌کند.

به طور کلی، تیم‌ها ساختار پروتئین‌ها را امسال با دقت بیشتری در مقایسه با آخرین CASP پیش بینی کردند اما اکثر این پیشرفت‌ها را آلفا فولد به دست آورده است. در مورد اهداف پروتئینی که نسبتاً دشوار به نظر می‌رسند، عملکرد سایر تیم‌ها در بهترین حالت 75 نمره از 100 نمره بود، در حالی که این عدد برای آلفا فولد 90 بود.



به گفته مولت تقریباً نیمی از تیم‌ها استفاده از «یادگیری عمیق» را در مقاله خود ذکر کرده‌اند که نشانگر تاثیر گسترده هوش مصنوعی در این زمینه علمی است. بیشتر شرکت کنندگان از تیم‌های دانشگاهی بودند اما مایکروسافت و شرکت چینی «تنسنت» نیز در CASP14 شرکت کرده بودند.

برخی از شرکت‌کنندگان CASP از جمله «محمد القریشی»، زیست‌شناس محاسباتی در دانشگاه کلمبیا در نیویورک مشتاق است که جزئیات عملکرد آلفا فولد را پس از ارائه این رورکدر توسط دیپ‌ماینر بررسی کند. القریشی معتقد است که نتایج تحقیقات آلفا فولد تحول آفرین خواهد بود. این متخصص می‌گوید:

«فکر می‌کنم منصفانه است که بگوییم این امر در زمینه پیش بینی ساختار پروتئین‌ها با روش‌های تجربی (مانند کریستالوگرافی) تاثیر منفی دارد. گمان می‌کنم بسیاری از دانشمندان

این روش‌های سنتی را رها می‌کنند چون بدون شک مشکل اصلی حل شده. مطمئناً این اتفاق یکی از مهمترین نتایج علمی زندگی من است.»

ساختارهای سریعتر

یکی از پیش بینی‌های تیم آلفا فولد به تعیین ساختار یک پروتئین باکتریایی کمک کرده که آزمایشگاه لوپاس سال هاست سعی در تعیین ساختار آن دارد. تیم لوپاس پیش از این داده‌های انعکاسی اشعه ایکس خام را جمع آوری کرده بود اما تبدیل این الگوها به یک ساختار پروتئینی به اطلاعاتی در مورد شکل پروتئین نیاز داشت. تلاش‌ها برای بدست آوردن این اطلاعات و سایر ابزارهای پیش بینی با شکست مواجه شده بود. لوپاس می‌گوید:

«مدل گروه 427 بعد از گذشت یک دهه که همه چیز را امتحان کردیم، تنها در نیم ساعت ساختاری را به ما داد که ده سال در جستجوی آن بودیم.»

«دمیس هساییس»، بنیانگذار و مدیر عامل دیپ‌ماینده، می‌گوید این شرکت قصد دارد آلفافولد را در مسیری قرار دهد تا دانشمندان دیگر هم امکان استفاده از آن را داشته باشند. هساییس، کشف دارو و طراحی پروتئین را از اهداف آلفافولد عنوان می‌کند:

«ما در حال درک بهتر خواسته‌های زیست‌شناسان هستیم که یکی کشف داروهای جدید و دیگری طراحی پروتئین هاست.»



دمیس هسابیس، مدیر عامل دیپ‌ماین

در اوایل سال 2020، این شرکت پیش بینی ساختارهایی از پروتئین‌های ویروس کرونا را که هنوز در آزمایش‌های تجربی اثبات نشده بودند، منتشر کرد. پیش بینی دیپ‌ماین در مورد پروتئینی به نام «Orf3a» در نهایت بسیار شبیه به پروتئینی است که بعداً از طریق روش Cryo-EM تعیین شد و این قدرت پیش بینی برنامه‌های آلفا فولد را نشان می‌دهد.

تاثیرات در دنیای واقعی

بعید است آلفا فولد به زودی بتواند درهای آزمایشگاه‌های عظیمی مانند «بروهاون» را ببندد که از روش‌های تجربی برای تحلیل ساختارهای پروتئینی استفاده می‌کنند. با این حال دستاورد آنها بدین معنی است که از داده‌های آزمایشی با کیفیت پایین‌تر که به شکلی آسان‌تر جمع‌آوری شده می‌توان برای تحلیل ساختار پروتئین‌ها بهره برد. برخی از برنامه‌های کاربردی، مانند تجزیه و تحلیل تکاملی پروتئین‌ها، قرار است شکوفا شوند زیرا سونامی داده‌های ژنومی موجود ممکن است برای تحلیل ساختاری بسیاری از پروتئین‌ها به کار رود. لوپاس می‌گوید:

«این کار به نسل جدید زیست‌شناسان مولکولی امکان می‌دهد به دنبال پاسخ برای پرسش‌های پیشرفته‌تری باشند.»

«جانت تورنتون»، زیست شناس ساختاری در «آزمایشگاه زیست شناسی مولکولی اروپا»، «انستیتوی بیوانفورماتیک اروپا» در انگلستان، و یکی از داوران پیشین CASP می‌گوید: «این مشکلی بود که فکر می‌کردم در طول زندگی من حل نخواهد شد.»

نتایج الگوریتم آلفا فولد آنقدر عظیم است که از نظر متخصصان همه جوانب زندگی را تغییر می‌دهد

این گفته جانت تورنتون به ما نشان می‌دهد چه اتفاق و چه انقلابی در این زمینه روی داده است. او امیدوار است این روش بتواند به روشن شدن عملکرد هزاران پروتئین ناشناخته در ژنوم انسان کمک کند و تغییرات ژنی ایجاد کننده بسیاری از بیماری‌ها را، که ممکن است در افراد متفاوت باشد، کشف کند.

عملکرد آلفا فولد جدای از دنیای زیست‌شناسی برای آینده دیپ‌مایند گوگل نیز نقطه عطفی است. این شرکت بیشتر به خاطر داشتن هوش مصنوعی در بازی‌هایی مانند Go مشهور شد، اما هدف طولانی مدت آن توسعه برنامه‌هایی با توانایی دستیابی به هوش گسترده و شبه انسانی است. هسایس در اینباره می‌گوید:

«مقابله با چالش‌های بزرگ علمی، مانند پیش بینی ساختار سه بعدی پروتئین، یکی از مهمترین کاربردهایی است که هوش مصنوعی می‌تواند انجام دهد. من فکر می‌کنم این مهمترین کاری است که ما از نظر تأثیر در دنیای واقعی انجام داده‌ایم.»

همانطور که مدیرعامل دیپ‌مایند می‌گویند حالا دیگر دوران نمایش قدرت هوش مصنوعی از بازی‌ها فراتر رفته و نوبت به بهره‌گیری از آن برای غلبه بر چالش‌های بزرگتر رسیده است. دیپ‌مایند تنها در چند سال توانست مساله‌ای ۵۰ ساله را برای متخصصان حل کند و باید دید در دهه آینده چه مشکلات عظیمی را از سر راه برخواهد داشت.

[دیجیاتو](#)