



امید به درمان همه بیماری‌ها به دست هوش مصنوعی

پایگاه داده ساختار پروتئین شرکت "الفبت" موسوم به "آلفا فولد" اکنون شامل بیش از ۲۰۰ میلیون پروتئین شناخته شده است که امیدها را برای درمان انواع بیماری‌ها زنده نگه می‌دارد.

پایگاه داده ساختار پروتئین شرکت "الفبت" موسوم به "آلفا فولد" اکنون شامل بیش از ۲۰۰ میلیون پروتئین شناخته شده است که امیدها را برای درمان انواع بیماری‌ها زنده نگه می‌دارد.

به گزارش ایسنا و به نقل از نیو اطلس، سال گذشته، هوش مصنوعی شرکت "الفبت" (Alphabet) به نام "دیپ مایند" (DeepMind)، یک پایگاه داده منبع باز از ساختارهای سه بعدی صدها هزار پروتئین از جمله تمام ۲۰ هزار پروتئین شناخته شده در بدن انسان منتشر کرد. اکنون این پایگاه داده موسوم به "آلفا فولد" به ۲۰۰ میلیون نمونه گسترش یافته است که تقریباً می‌توان گفت تمام پروتئین‌های شناخته شده در دنیای علم را شامل می‌شود.

پروتئین‌ها عامل فعالیت سلول‌های زنده هستند که تعداد بی‌شماری از فرآیندهای بیولوژیکی حیاتی در راه حیات را انجام می‌دهند. آنها از زنجیره‌ای از اسیدهای آمینه تشکیل شده‌اند که به شکل‌های سه بعدی پیچیده می‌شوند که عملکرد آنها را دیکته می‌کند. ترسیم ساختار پروتئین‌ها برای درک اینکه چه کاری انجام می‌دهند، چگونه کار می‌کنند و اینکه چگونه ممکن است چیزی اشتباه پیش برود، مهم است. شناخت و ترسیم پروتئین‌ها برای تحقیق در مورد همه چیز، از داروها و درمان‌های جدید گرفته تا بهبود محصولات کشاورزی و حفاظت از حیوانات، کلیدی است.

اما محاسبه ساختار دقیق یک پروتئین بر اساس اسیدهای آمینه تشکیل دهنده آن دشوار است. پی بردن به این موضوع معمولاً به مقدار زیادی قدرت محاسباتی و ساعات کار سرسام‌آور انسانی نیاز دارد، در حالی که این وضعیت به «مشکل تاخوردگی پروتئین» معروف شده است. به این ترتیب، پیشرفت در این حوزه در طول دهه‌های اخیر نسبتاً کند بوده است.

اکنون می‌توان گفت این مشکل تا زمانی مشکل بود که شرکت "الفبت" که شرکت مادر "گوگل" است، هنوز هوش مصنوعی قدرتمند خود موسوم به "دیپ مایند" را روی این مشکل تنظیم نکرده بود. این سیستم که در ابتدا با ۱۰۰ هزار ساختار پروتئینی شناخته شده آموزش دیده بود، توانایی پیش‌بینی ساختار میلیون‌ها پروتئین دیگر را توسعه داد که تعیین و ترسیم هر یک از آنها به جای ماه‌ها یا سال‌ها، تنها چند دقیقه یا چند ثانیه زمان می‌برد.

در ژوئیه ۲۰۲۱ اولین پایگاه داده ساختار پروتئین "آلفا فولد" برای عموم منتشر شد. این پایگاه داده در ابتدا حاوی بیش از ۲۵۰ هزار ساختار پروتئینی بود که حدود ۹۸.۵ درصد از پروتئین‌های انسانی و همچنین پروتئین‌های موجود در مگس‌های میوه، موش‌ها، مخمرها و باکتری‌های "ای.کولی" (E.coli) را شامل می‌شد. بعدها به حدود یک میلیون ساختار پروتئینی از ۱۰ هزار گونه جانوری، گیاه، باکتری، قارچ و سایر موجودات گسترش یافت.

حالا ظرف یک سال پس از آن رویداد، بیش از ۵۰۰ هزار دانشمند از سراسر جهان برای کمک گرفتن در تحقیقات خود به این پایگاه داده دسترسی پیدا کرده‌اند.

اکنون "دیپ مایند" یک به روزرسانی بزرگ جدید برای پایگاه داده خود منتشر کرده است که در حال حاضر شامل حدود ۲۱۴ میلیون ساختار از یک میلیون گونه است. این عدد یعنی تقریباً تمام پروتئین‌هایی را که در حال حاضر برای علم شناخته شده است، پوشش می‌دهد و برای تحقیق در مورد درمان بیماری‌ها، واکسن‌ها، پایداری، مقاومت آنتی‌بیوتیکی و حتی آلودگی پلاستیکی موهبت بزرگی را ارائه می‌دهد.

"اریک توپول"، مدیر مؤسسه تحقیقاتی "اسکریپس" می‌گوید: "آلفا فولد" قبلاً اکتشافات عظیمی از جمله شکستن ساختار مجتمع منافذ هسته‌ای را تسریع و فعال کرده است و با افزودن این ساختارهای جدید که تقریباً کل جهان پروتئین‌ها را روشن می‌کند، می‌توانیم انتظار داشته باشیم که اسرار بیولوژیکی بیشتری هر روز رمزگشایی شود.

گفتنی است که کل پایگاه داده ساختارهای پروتئین، متشکل از بیش از ۲۵ ترابایت داده را می‌توان از مجموعه داده‌های عمومی Google Cloud دانلود کرد.