

۱۰ کد رایانه‌ای که علم را متحول کردند

بدون برنامه‌نویسی و کدهای رایانه‌ای انجام بسیاری از کارهای علمی امکان پذیر نمی‌شد و جهان هیچ‌گاه به پیشرفت‌های امروزی نمی‌رسید. اما ۱۰ کدی که جهان را تغییر دادند چه بودند؟



بدون برنامه نویسی و کدهای رایانه ای انجام بسیاری از کارهای علمی امکان پذیر نمی شد و جهان هیچ گاه به پیشرفت های امروزی نمی رسید. اما ۱۰ کدی که جهان را تغییر دادند چه بودند؟

به گزارش ایسنا و به نقل از نیچر، در سال ۲۰۱۹ گروه تحقیقاتی "تلسکوپ افق رویداد" (Event Horizon Telescope) اولین تصویر از سیاه چاله را به جهانیان نشان دادند. اما تصویر درخشان یک شی حلقه مانند که گروه از آن رونمایی کردند یک تصویر عادی نبود بلکه یک تصویر رایانه ای بود.

با انجام محاسبات ریاضیاتی بر روی داده های به دست آمده از تلسکوپ های رادیویی در ایالات متحده، مکزیک، شیلی، اسپانیا و قطب جنوب این داده ها تغییر کرده و تبدیل به تصویر شد. این گروه کد کامپیوتری که برای انجام این کار استفاده کرده بودند را در کنار مقالاتشان منتشر کردند تا جامعه ی علمی کاری که آن ها انجام داده بودند را ببیند و از آن برای رسیدن به موفقیت های بعدی استفاده کنند. این یک الگوی رو به افزایش است. از ستاره شناسی گرفته تا جانورشناسی پشت هر کشف بزرگ علمی در عصر جدید یک کامپیوتر وجود دارد.

مایکل لویت (Michael Levitt) یک زیست شناس محاسباتی در دانشگاه "استنفورد" کالیفرنیا یکی از برندگان جایزه ی نوبل شیمی در سال ۲۰۱۳ بوده است. او این جایزه را برای کار بر روی استراتژی های مدل سازی در ساختارهای شیمیایی دریافت کرد. لویت خاطرنشان کرد که لپ تاپ های امروزی ۱۰ هزار برابر حافظه و سرعت بیشتری نسبت به کامپیوتر آزمایشگاه او در سال ۱۹۶۷ دارند. زمانی که کار بر روی پروژه اش که برنده ی نوبل شد را آغاز کرد.

او می گوید: ما امروزه تجهیزات محاسباتی فوق العاده ای در دست داریم اما مشکل اینجاست که هنوز نیاز به فکر کردن داریم. یک کامپیوتر قوی بدون وجود نرم افزاری که بتواند سوالات دانشمندان را پاسخ دهد و محققانی که بدانند چگونه کدنویسی کنند و از آن استفاده کنند بی مصرف است.

نیل چو هونگ (Neil Chue Hong) مدیر موسسه ی "Software Sustainability" که دفتر مرکزی آن در ادینبرو اسکاتلند است و مرکزی برای بهبود توسعه و استفاده از نرم افزار در علوم است می گوید: تحقیقات امروزه کاملاً به نرم افزارها وابسته هستند. و پارامترهای آن تمام جنبه های تحقیقات را شامل می شود.

کشف های جدید در رسانه ها منتشر می شوند اما در اینجا می خواهیم پشت صحنه ی این کشف ها و بخش اصلی آن یعنی کدهایی که تحقیقات را در طی چند دهه ی گذشته متحول کرده اند بررسی کنیم.

اگرچه هیچ لیست این چنینی نمی تواند قطعی باشد اما ما طی سال گذشته از ده ها محقق نظرسنجی کردیم تا مجموعه ای از ۱۰ نرم افزار که تأثیری بزرگ بر علم داشتند را مشخص کنیم.

همگردان فورترن: پیشگام در زبان برنامه نویسی (۱۹۵۷)

استفاده از اولین کامپیوترها کار راحتی نبود و برنامه نویسی به معنای واقعی کلمه با دست و با اتصال مجموعه ی مدارها و سیم ها انجام می شد. دستگاه هایی که بعد از آن آمد و زبان برنامه نویسی هم گذاری (نوعی برنامه نویسی سطح پایین) به کاربران کمک کرد تا کامپیوترها را با کد برنامه نویسی کنند. اما همچنان نیاز به دانش ساختاری کامپیوتر بود. بنابراین هنوز این کار از دسترس بیشتر دانشمندان خارج بود.

این موضوع در دهه ی ۱۹۵۰ با توسعه ی زبان برنامه نویسی نمادین به خصوص با توسعه ی زبان برنامه نویسی فورترن (Fortran) که توسط جان باکس (John Backus) و گروهش در موسسه ی آی بی ام در سن خوزه کالیفرنیا ایجاد شده بود تغییر کرد. با استفاده از زبان برنامه نویسی فورترن کاربران می توانند برنامه نویسی را با دستورالعمل هایی که برای انسان ها قابل خواندن است انجام دهند. یک همگردان (compiler) سپس این دستورالعمل ها را به یک کد سریع و کارآمد ماشینی تبدیل می کند.

اگرچه هنوز هم برنامه نویسی کار آسانی نبود. در ابتدای کار برنامه نویسان از مجموعه ای کارت برای وارد کردن کدها استفاده می کردند و برای یک شبیه سازی پیچیده ده ها هزار کارت نیاز بود. سیوکورو مانابه (Syukuro Manabe) اقلیم شناس در دانشگاه پرینستون نیوجرسی می گوید: فورترن برنامه نویسی را برای آن دسته از محققانی که دانشمند در زمینه کامپیوتر نبودند نیز ممکن کرد. برای اولین بار توانستیم خودمان برنامه نویسی را انجام دهیم. او و همکارانش از این زبان برنامه نویسی استفاده کردند تا یکی از موفق ترین مدل های اقلیمی را بسازند.

اکنون پس از گذشت هشت دهه فورترن همچنان به طور گسترده ای استفاده می شود. از جمله در مدل سازی اقلیمی، دینامیک سیالات، شیمی محاسباتی و هر موضوعی که در آن از جبر خطی پیچیده استفاده شود و نیاز به کامپیوترهای قدرتمند برای بررسی سریع اعداد باشد. کدی که در نتیجه به دست می آید سریع است و هنوز تعداد زیادی از برنامه نویسان وجود دارند که می دانند چگونه آن را بنویسند. نوع قدیمی برنامه نویسی فورترن هنوز هم در آزمایشگاه ها و ابررایانه ها در سراسر جهان مورد استفاده قرار می گیرد.

پردازنده سیگنال: "تبدیل فوریه" سریع (۱۹۶۵)

هنگامی که اخترشناسان رادیویی آسمان را بررسی می کردند سیگنالی پیچیده یافتند که با گذر زمان تغییر می کرد. برای فهمیدن ماهیت این امواج رادیویی آن ها باید می فهمیدند که این سیگنال ها به عنوان تابع فرکانسی چه شکلی دارند. انجام یک فرآیند ریاضیاتی به نام "تبدیل فوریه" (Fourier transform) به محققان در انجام این کار کمک می کرد. مشکل این بود که این روش ناکارآمد بود و برای یک مجموعه داده با پایه N به محاسباتی با پایه N به توان ۲ نیاز است.

در سال ۱۹۶۵ ریاضی دانان ایالات متحده، جیمز کولی (James Cooley) و جان توکی (John Tukey) راه حلی برای بهبود این فرآیند پیدا کردند. آن ها با استفاده از روش "بازگشت" رویکرد برنامه نویسی "تقسیم و حل" را که در آن الگوریتم مدام خودش را تکرار می کند، فرآیند برنامه نویسی "تبدیل فوریه" سریع را آسان تر کرد که با افزایش N سرعت نیز افزایش پیدا می کرد و به طور مثال برای عدد ۱۰۰۰ سرعت ۱۰۰ برابر بیشتر می شود.

نیک ترفتن (Nick Trefethen)، ریاضیدان دانشگاه آکسفورد می گوید: این کشف در اصل کشفی بود که قبلا اتفاق افتاده بود توسط ریاضیدان آلمانی کارل فردریش گاس (Carl Friedrich Gauss) در سال ۱۸۰۵ اما او هیچوقت آن را منتشر نکرد اما کولی و توکی این کار را انجام دادند و این یک اتفاق فوق العاده در ریاضیات کاربردی و مهندسی است. از کد "تبدیل فوریه" بارها و بارها استفاده شده است.

پاول آدامز (Paul Adams)، مدیر بخش بیوفیزیک مولکولی و تصویربرداری بیولوژیکی در آزمایشگاه ملی لارنس برکلی (Lawrence Berkeley National Laboratory) است. او می گوید هنگامی که در سال ۱۹۹۵ ساختار پروتئین باکتریایی GroEL را اصلاح و پالایش می کردیم این محاسبات به کمک فرآیند "تبدیل فوریه" ساعت ها طول کشید و اگر کمک این فرآیند نبود نمی دانم چگونه این کار را می کردیم. احتمالا تا ابد به طول می انجامید.

دسته بندی مولکولی: بانک داده های بیولوژیکی (۱۹۶۵)

بانک های اطلاعاتی مولفه های یکپارچه ی تحقیقات علمی امروزه هستند که به راحتی می توان از این حقیقت که آن ها به وسیله ی نرم افزارها تهیه می شوند چشم پوشی کرد. در طول چند دهه ی گذشته این منابع از نظر اندازه بسیار رشد کرده اند و بسیاری از حوزه ها را شکل داده اند اما این تغییرات در هیچ کجا به اندازه ی حوزه ی زیست شناسی قابل توجه نبوده است. مجموعه داده های عظیمی که امروزه از ژنوم ها و پروتئین ها داریم ریشه در کارهای مارگارت دیهوف (Margaret Dayhoff) دارد. او پیشگام در بیوانفورماتیک در بنیاد تحقیقات ملی پزشکی در سیلور اسپرینگ (Silver Spring) مریلند (Maryland) بود. در اوایل دهه ی ۱۹۶۰، هنگامی که زیست شناسان تلاش می کردند تا توالی آمینواسید را جداسازی کنند، دیهوف شروع به جمع آوری این اطلاعات برای جستجوی سرخ هایی از روابط تکاملی بین گونه های مختلف کرد.

اطلس او از توالی پروتئین ها و ساختار آن ها در سال ۱۹۶۵ با همکاری سه نویسنده ی دیگر منتشر شد. آن ها در این مقاله آنچه را که به عنوان توالی، ساختار و شباهت های ۶۵ پروتئین شناخته می شد توضیح دادند. برونو استریسر (Bruno Strasser) که یک مورخ است در سال ۲۰۱۰ نوشت: این مجموعه اولین مجموعه ای بود که برای پاسخ به یک سوال در یک تحقیق خاص مربوط نبود. کدهای آن بر روی کارت هایی پانچ شدند تا بتوان این مجموعه داده را گسترش داد و جستجو کرد.

"بانک داده پروتئین" که امروزه جزئیات بیش از ۱۷۰ هزار ساختار درشت مولکول را در اختیار دارد کار خود را در سال ۱۹۷۱ آغاز کرد. راسل دولیتل (Russell Doolittle) یک زیست شناس تکاملی از دانشگاه کالیفرنیا در سن دیگو مجموعه داده پروتئین دیگری به نام "Newat" در سال ۱۹۸۱ ایجاد کرد و در سال ۱۹۸۲ بانک داده دیگری به نام "بانک ژن" (GenBank) ایجاد شد این مجموعه شامل داده های دی ان ای (DNA) است که توسط موسسه ملی بهداشت ایالات متحده نگهداری می شود.

ارزش چنین منابعی در ژوئیه سال ۱۹۸۳ مشخص شد هنگامی که گروه های مجزا به رهبری مایکل واترفیلد (Michael Waterfield) یک زیست شناس پروتئین و شیمی دان در موسسه ی تحقیقات سرطان ایمپریال در لندن (Imperial Cancer Research Fund in London) و دولیتل به طور مستقل شباهت هایی در توالی یک عامل خاص در رشد انسان و پروتئینی در ویروس ایجاد کننده ی سرطان در میمون ها گزارش دادند. این مشاهدات حاکی از مکانیزمی در ویروس بود که عامل رشد را شبیه سازی می کرد و باعث رشد کنترل نشده ی سلول ها می شد.

جیمز استل (James Ostell)، مدیر سابق مرکز ملی اطلاعات زیست فناوری ایالات متحده می گوید: برای آن دسته از زیست شناسان که به کامپیوتر و آمار علاقه مند نیستند باید بگویم ما می توانیم با بررسی توالی ها اطلاعاتی درباره ی سرطان ها به دست آوریم و علاوه بر طراحی آزمایش ها برای بررسی فرضیه ها، محققان می توانند داده های عمومی به دست آورند که تاکنون افرادی که واقعا داده ها را جمع آوری می کنند به آن دست نیافته باشند و هنگامی که داده های مختلف به یکدیگر متصل شوند این قدرت افزایش می یابد. موضوعی که محققان مرکز ملی اطلاعات زیست فناوری (NCBI) در سال ۱۹۹۱ به وسیله ی جستجوگر پیشرفته ی آنتره (Entrez) به آن دست یافتند. این جستجوگر به محققان کمک می کند تا آزادانه از موضوعی همچون DNA به موضوع ادبیات یا پروتئین بروند و بازگردند.

پیش بینی آب و هوا: مدل سازی دورانی (۱۹۶۹)

با پایان جنگ جهانی دوم، جان ون نیومن (John von Neumann) پیشگام در عرصه ی کامپیوتر، کامپیوترهایی کرد که پیش از آن برای محاسبه ی مسیر پرواز و طراحی اسلحه استفاده می شدند برای حل مشکل پیش بینی آب و هوا تغییر کاربری داد.

مانابه توضیح می دهد: تا آن زمان پیش بینی آب و هوا تنها کاری تجربی بود که با استفاده از تجربه و ظن و گمان، وضعیت جوی آینده را پیش بینی می کردند. اما گروه ون نیومن قصد داشتند تا با استفاده از قوانین فیزیک به صورت عددی آب و هوا را پیش

بینی کنند.

ونکاترامانی بالاجی (Venkatramani Balaji)، رئیس بخش مدل سازی سیستم ها در آزمایشگاه دینامیک سیالات ژئوفیزیکی اداره ملی اقیانوس شناسی و اقلیمی پرینستون در نیوجرسی می گوید: این معادلات برای دهه ها شناخته شده بودند اما هواشناسان اولیه نتوانستند آن ها را حل کنند. برای انجام این کار باید شرایط جوی موجود در آن لحظه وارد می شد و تغییرات آن در طول مدتی کوتاه بررسی می شد و این کار تکرار می شد. این فرآیند بسیار طولانی و وقت گیر بود و تکمیل معادله ی ریاضیاتی آن پیش از تغییر آب و هوا ممکن نبود. در سال ۱۹۲۲ لوئیس فرای ریچاردسون (Lewis Fry Richardson)، یک ریاضیدان ماه ها صرف انجام محاسباتی کرد تا آب و هوای مونیخ آلمان را پیش بینی کند اما نتایج آن کاملاً غلط بودند از جمله آن پیش بینی هایی بود که تحت هیچ شرایطی رخ نخواهند داد. کامپیوترها این مشکل را قابل حل کردند.

در اواخر دهه ی ۱۹۴۰ ون نیومن گروه خود را برای پیش بینی وضعیت آب و هوا در موسسه ی مطالعات پیشرفته پرینستون تاسیس کرد. در سال ۱۹۵۵ گروه دیگری کار بر روی "پیش بینی بی نهایت آب و هوا" را آغاز کردند که نوعی مدل سازی جوی است.

مانابه که در سال ۱۹۵۸ به گروه مدل سازی جوی پیوست قصد داشت بر روی مدل های جوی کار کند. همکار او کریک برایان (Kirk Bryan) نیز بر روی اقیانوس ها کار می کرد. در سال ۱۹۶۹ آن ها موفق شدند تا کارهایشان را ترکیب کنند و آنچه را که مجله ی نیچر در سال ۲۰۰۶ "نقطه عطفی" در محاسبات علمی خواند ایجاد کنند.

مدل های امروزی می توانند سطح سیاره را به مربع های ۲۵ در ۲۵ کیلومتری تقسیم کنند و جو را به ده ها سطح تقسیم می کند. مدل ترکیبی مانابه و برایان از مربع های ۵۰۰ کیلومتری و ۹ سطح استفاده می کرد و تنها یک ششم کره ی زمین را پوشش می داد اما بالاجی می گوید: هنوز هم آن مدل کار فوق العاده ای انجام داد و گروه توانست برای اولین بار تاثیر افزایش سطح دی اکسید کربن را به وسیله ی فرآیند "درون رایانه ای" بررسی کند.

حسابگر بلاس (۱۹۷۹) (BLAS)

محاسبات علمی معمولاً شامل ریاضیات ساده و نمودار و ماتریس است فقط تعداد آن ها زیاد است اما در دهه ی ۱۹۷۰ هیچ ابزار جهانی برای انجام این محاسبات وجود نداشت در نتیجه برنامه نویسانی که در حوزه علم کار می کردند زمانشان را صرف تولید کدهایی کارآمد برای انجام محاسبات ریاضیاتی می کردند به جای آن که بر روی سوالات علمی تمرکز کنند.

جهان برنامه نویسی نیاز به یک ابزار استاندارد داشت و در سال ۱۹۷۹ به آن دست یافت: زیربرنامه ی جبر خطی پایه BLAS، استاندارد است که تا سال ۱۹۹۰ به تکامل خود ادامه داد و ده ها جریان اساسی برای نمودارها و ماتریس ها مشخص کرد.

جک دونگارا (Jack Dongarra)، دانشمند رایانه در دانشگاه تنسی (Tennessee) که عضو گروه توسعه ی بلاس بود می گوید: در واقع BLAS محاسبات ریاضی ماتریس و نمودار را به یک واحد محاسباتی کاهش می دهد.

تصاویر میکروسکوپی سازمان بهداشت جهانی (۱۹۸۷)

در اوایل دهه ی ۱۹۸۰، وین رازبند (Wayne Rasband) در حال کار با یک آزمایشگاه تصویربرداری از مغز در موسسه ی ملی بهداشت ایالات متحده در بتسدا (Bethesda) مرلند بود. گروه یک اسکنر برای رقومی کردن (دیجیتالی کردن) (digitize) فیلم های اشعه ایکس داشت اما راهی برای تحلیل آن ها در کامپیوتر وجود نداشت. بنابراین رازبند برنامه کامپیوتری برای این موضوع نوشت. این برنامه برای کامپیوتر کوچک غیر شخصی ۱۵۰ هزار دلاری آمریکا به نام "PDP-۱۱" نوشته شده بود. سپس در سال ۱۹۸۷ شرکت اپل از کامپیوتر شخصی اش به نام "Macintosh II" رونمایی کرد که استفاده از آن راحت تر بود و مقرون به صرفه تر نیز بود.

رازبند می گوید: این کامپیوتر جدید به نظر من به عنوان یک آزمایشگاه بررسی تصویر بسیار بهتر کار می کرد. او نرم افزارش را به یک پلتفرم جدید منتقل کرد و نام آن را نیز تغییر داد. سیستم تصویر برداری موسسه ملی بهداشت به محققان کمک کرد تا هر تصویری را درون هر رایانه ای مشاهده کنند. این نرم افزار شامل ایمج جی (ImageJ) نرم افزاری مبتنی بر جاوا (JAVA) بود که رازبند آن را برای کاربران ویندوز و سیستم عامل لینوکس (Linux) نوشت.

بث سیمینی (Beth Cimini)، یک زیست شناس محاسباتی که بر روی پلتفرم تصویربرداری موسسه ی کمبریج ماساچوست کار می کند می گوید: ایمج جی در حال حاضر اساسی ترین ابزاری است که ما در اختیار داریم، من هیچ زیست شناسی نمی شناسم که از میکروسکوپ استفاده کرده باشد اما از ایمج جی استفاده نکرده باشد.

رازبند می گوید: یکی از علت های آن رایگان بودن این برنامه است. کوین الیسیری (Kevin Eliceiri) مهندس پزشکی دانشگاه ویسکانسین (Wisconsin) می گوید: همچنین آسان بودن این برنامه و امکان تنظیم ابزار آن بسته به نیاز این ابزار را برای کاربران مناسب کرده است.

جستجوگر توالی: بلاست (۱۹۹۰)

برای یک نرم افزار هیچ چیز بهتر از آن نیست که اسمش تبدیل به یک فعل در زبان شود. برای جستجو، گوگل و برای ژنتیک، "بلاست" (BLAST).

تغییرات تکاملی را می توان در توالی های مولکولی یافت. مانند جابه جایی، حذف، فاصله و بازسازی. با جستجوی شباهت های بین این توالی ها مخصوصاً در میان پروتئین ها، محققان می توانند ارتباطات تکاملی را کشف کنند و بینشی در مورد عملکرد ژن ها به دست آورند. برای این کار داده های مولکولی که به سرعت رشد می کنند باید سریع و جامع بررسی شوند.

دیپهوف در سال ۱۹۷۸ بخش مهمی از این مشکل را حل کرد. او ماتریس "جهش پذیرفته نقطه ای" ابداع کرد که به محققان اجازه

می داد تا ارتباط دو پروتئین را نه تنها براساس شباهت توالی بلکه براساس فاصله تکاملی آن ها ارزیابی کنند. در سال ۱۹۸۵ ویلیام پیرسون (William Pearson) از دانشگاه ویرجینیا شارلوتسویل (Virginia in Charlottesville) و دیوید لیپمن (David Lipman) از مرکز ملی اطلاعات زیست فناوری الگوریتمی به نام "FASTP" را معرفی کردند که ماتریس دیهوف را با توانایی جستجوی سریع ادغام می کرد.

سال ها بعد لیپمن به همراه وارن گیش (Warren Gish) و استفن آلتچول (Stephen Altschul) از مرکز ملی اطلاعات زیست فناوری، وب میلر (Webb Miller) از دانشگاه ایلاتی پنسیلوانیا و جین میرز (Gene Myers) از دانشگاه آریزونا روش قدرتمندتری از پالایش را ابداع کردند: ابزار پایه ای برای جستجوی برهم نهی های موضعی یا "بلاست" بلاست که در سال ۱۹۹۰ منتشر شد سرعت جستجو را به حدی رساند که از پس رشد داده ها بریاید و همچنین شباهت های داده هایی که از نظر تکاملی فاصله داشتند پیدا کند. این ابزار همزمان می توانست احتمال رخ دادن چنین شباهت هایی را بررسی کند.

آلتچول می گوید: نتایج به دست آمده سریع بودند و مهم تر اینکه استفاده از آن آسان بود. در دوره ای که اطلاعات به وسیله ی پست منتقل می شدند گیش یک سیستم پست الکترونیکی و بعد از آن هم ساختاری مبتنی بر وب ایجاد کرد تا محققان بتوانند کامپیوترهای مرکز ملی اطلاعات زیست فناوری را از راه دور جستجو کنند تا مطمئن شوند داده هایی که در دست دارند به روز است.

منبع مقالات پیش چاپ (۱۹۹۱) arXiv.org

در اواخر دهه ی ۹۰ فیزیکدانان ذرات به طور معمول نسخه هایی از کار خود را برای همکارانشان پست می کردند تا نظر آن ها را بدانند.

در سال ۱۹۹۱، پائول گینسپراگ (Paul Ginsparg) یک فیزیکدان، که در آن زمان در آزمایشگاه ملی آلاموس مکزیک کار می کرد یک ایمیل خودکار ایجاد کرد که به صورت خودکار ارسال می شد تا زمینه ای برابر برای همه ی افراد در زمینه دسترسی به مقالات ایجاد کند. مشترکین روزانه لیستی از مقالات پیش از چاپ دریافت می کردند که هرکدام شناسه مخصوص خود را داشتند و با یک ایمیل ساده کاربران می توانستند مقاله ای را از سیستم کامپیوتری آزمایشگاه دریافت یا بازیابی کنند و همچنین لیستی از مقالات جدید بیابند و یا براساس نام مقاله یا نام نویسنده آن را جستجو کنند.

گینسپراگ قصد داشت تا مقالات را برای مدت سه ماه حفظ کند و سپس محتوی را محدود به جامعه ی فیزیک ذرات کند اما یکی از همکارانش او را متقاعد کرد که مقالات را به طور بی نهایت حفظ کند.

در سال ۱۹۹۳، گینسپراگ این سیستم را به شبکه جهانی وب منتقل کرد و در سال ۱۹۹۸ نامی به آن داد که تا امروز باقی مانده است: arXiv.org

این منبع اطلاعاتی که اکنون سی امین سال خود را می گذراند شامل ۱.۸ میلیون پیش چاپ مقالات به طور رایگان است و ماهانه ۱۵ هزار مقاله ارسالی و ۳۰ میلیون دانلود دارد.

جستجوگر داده ها: نوت بوک آی پایتون (۲۰۱۱)

فرناندو پرز (Fernando Perez) که دانشجوی کارشناسی ارشد در سال ۲۰۰۱ بود، تصمیم گرفت مولفه های اصلی زبان برنامه نویسی پایتون (Python) را در دست بگیرد.

پایتون یک زبان برنامه نویسی تفسیری است که یعنی برنامه ها خط به خط اجرا می شوند. برنامه نویسان می توانند از ابزاری به نام "REPL" استفاده کنند و کدی بنویسند که برنامه ای به نام مفسر آن را اجرا کند. "REPL" به جستجوی سریع و تکرار را فراهم می کند اما پرز خاطرنشان کرد که پایتون برای علم ساخته نشده بود بنابراین او نسخه ی خودش را ساخت که نتیجه ی آن آی پایتون (IPython) شد. یک مفسر تعاملی که پرز در دسامبر ۲۰۰۱ از آن رونمایی کرد. یک دهه بعد پرز با همکاری برایان گرینجر (Brian Granger) یک فیزیکدان و ایوان پترسون (Evan Patterson) یک ریاضیدان این ابزار را به مرورگر انتقال دادند. با راه اندازی آی پایتون نوت بوک انقلابی در عرصه ی داده های علمی ایجاد شد. این نرم افزار مانند سایر نوت بوک ها کدها، نتایج و متن را در یک فایل ادغام می کرد اما برخلاف آنها یک نرم افزار باز بود و از توسعه دهندگان دعوت می کرد تا در توسعه آن سهمی داشته باشند.

الکس نت (۲۰۱۲) AlexNet

هوش مصنوعی دو نوع دارد. یکی از قوانین کدگذاری شده استفاده می کند و دیگری یک کامپیوتر را قادر می سازد تا با الگوبرداری از ساختار عصبی مغز "یادگیری" انجام دهد.

جفری هینتون (Geoffrey Hinton) دانشمند رایانه از دانشگاه تورنتو کانادا می گوید: برای دهه ها محققان حوزه ی هوش مصنوعی این رویکرد را "بی معنی" می خواندند و آن را رد می کردند. در سال ۲۰۱۲ دانشجویان هینتون، الکس کریژفسکی (Alex Krizhevsky) و ایلیا ساتسکور (Alex Krizhevsky) خلاف این موضوع را ثابت کردند.

هینتون می گوید: در همایشی سالانه که در ImageNet برگزار می شود رقابتی بین محققان ایجاد می شود تا یک هوش مصنوعی را به وسیله ی داده های یک میلیون تصویر از اجسام روزمره تربیت کنند و سپس الگوریتم به دست آمده را بر روی یک مجموعه عکس جداگانه آزمایش کنند. در زمانی که بهترین الگوریتم ها یک چهارم دسته بندی ها را اشتباه انجام می دادند الگوریتم "یادگیری عمیق" کریژفسکی و ساتسکور که مبتنی بر شبکه ی عصبی بود این خطا را تا ۱۶ درصد کاهش داد. هینتون می گوید: ما اساسا میزان خطا را نصف کردیم.

هینتون می گوید: موفقیت گروه در سال ۲۰۱۲ نشان دهنده ی ترکیبی از مجموعه داده های آموزش، برنامه نویسی عالی و قدرت نوظهور واحد پردازش گرافیکی است. ناگهان ما موفق شدیم الگوریتم را ۳۰ برابر سریع تر اجرا کنیم. وی افزود: موفقیت واقعی در اصل سه سال قبل تر رخ داد، هنگامی که آزمایشگاهش شبکه ی عصبی ساخت که می توانست گفتار را با دقتی بیشتر از سایر هوش های مصنوعی تشخیص دهد. این موفقیت ها باعث شده که امروزه تلفن های همراه دستوره های گفتاری را بفهمند و ابزارهای تجزیه و تحلیل عکس بتوانند سلول ها را تشخیص دهند و به همین علت است که AlexNet در میان بسیاری از ابزاری که علم و جهان را متحول کرده اند جای می گیرد.