فرهنگ و اندیشهٔ ریاضی

سال ۴۰، شمارهٔ ۶۸ (بهار و تابستان ۱۴۰۰) صص. ۱۴۳ تا ۱۵۵

آغاز عصری جدید در ریاضیات کاربردی*

وِینان ای

مترجم: شاهرخ اسماعیلی

چکیده

در تاریخ علم، دو دورهٔ زمانی وجود داشته که بیشترین تأثیر را در ریاضی کاربردی ایجاد کرده است. زمان نیوتن که در طی آن، مشخص شد ریاضیات باید زبان علم باشد، و زمان فون‌نویمان که در آن، پیشنهاد شد الگوریتم‌های عددی باید پل اصلی بین ریاضیات و علم باشند. اکنون زمان سوم در پیش است؛ زمانی که اجزای اصلی ریاضی کاربردی شامل مدل‌سازی، روش‌های داده‌محور و الگوریتم‌ها طوری در جای خود قرار می‌گیرند که ریاضی کاربردی را بنیادی برای پژوهش علمی بین‌رشته‌ای و نوآوری فناورانهٔ هیجان‌انگیز معرفی کنند.

۱. الگوهای کپلری و نیوتنی

از زمان نیوتن، دو الگوی متفاوت برای انجام پژوهش علمی وجود داشته است: الگوی کپلری و الگوی نیوتنی. در الگوی کپلری، یا رویکرد داده‌محور[۱]، پژوهشگر از راه تحلیل داده‌ها، یافته‌های علمی را استخراج می‌کند. مثال کلاسیک در این زمینه، قوانین حرکت سیاره‌ایِ کپلر است. زیست‌داده‌ورزی[۲] نیز تصویری موجّه از موفقیت الگوی کپلری را در دوران نوین ارائه می‌دهد. در الگوی نیوتنی یا رویکرد مبتنی بر اصل اول، هدف کشف اصول بنیادی حاکم بر جهان پیرامون ما یا چیزهایی است که به آنها علاقه‌مندیم.

* نام و نشان مقاله به زبان اصلی از این قرار است:
Weinan E., The dawning of a new era in applied mathematics, *Notices*, **68** (2021), no. 4, 565–571.

[۱] پژوهشکدهٔ آمار برای معادل فارسیِ data-driven، «داده‌رهنمون» را پیشنهاد کرده است.

[۲] bioinformatics

بهترین مثال در این زمینه کارهای نیوتن، مَکسوِل، بولتسمان، اینشتین، هایزنبرگ و شرودینگر در فیزیک نظری است. امروزه فیزیک نظری هنوز هم یک زمین بازی مهم برای بهترین ذهن‌ها است. با پیشرفت روش‌های آماری و یادگیری ماشین، رویکرد داده‌محور به ابزاری بسیار توانا تبدیل شده است. هرچند این رویکرد برای یافتن واقعیت‌ها بسیار مؤثر است، درکمک به ما برای یافتن دلایل این واقعیت‌ها تأثیر کمی دارد.

رویکرد مبتنی بر اصل اول، رسیدن به فهم در بنیادی‌ترین سطح است. فیزیک، به‌ویژه، با دنباله‌روی از چنین اصول اولیه‌ای هدایت می‌شود. یک نقطهٔ عطف در این زمینه، پایه‌گذاری مکانیک کوانتمی در سال ۱۹۲۹ بود: همان‌طور که دیراک اعلام کرد [۲]، با وجود مکانیک کوانتومی، ما اکنون اصول اولیهٔ لازم برای بسیاری از مهندسی و علوم طبیعی بجز فیزیک در مقیاس استثنایی را در دست داریم.

با وجود این، همان‌طور که دیراک نیز اشاره کرد، مسئله ریاضیاتی که قوانین مکانیک کوانتومی را توصیف می‌کند بیش از اندازه پیچیده است. یکی از این دشواری‌ها، مسئله‌ای چندپیکره است که در آن با افزودن یک الکترون، ابعاد مسئله ۳ برابر می‌شود. این معضلی است که اغلب در رویکرد مبتنی بر اصل اول با آن مواجه می‌شویم؛ اصلی که بنیادی است اما چندان کاربردی نیست. در نتیجه در عمل، اغلب مجبور می‌شویم نظریه‌های دقیق و ظریف را کنار بگذاریم و به تقریب‌های موردی و بی‌قاعده متوسّل شویم. زیان این کار نه‌تنها نبود دقت و ظرافت است، بلکه اعتمادپذیری و انتقال‌پذیری نتایج را نیز از بین می‌برد.

ریاضی کاربردی در مسیری مشابه توسعه یافته است. چون نخستین اصول فیزیک برحسب معادله‌های دیفرانسیل با مشتقات جزئی (PDEها) صورت‌بندی شده‌اند، تحلیل و الگوریتم‌های عددی برای PDEها، به‌ویژه در طول دوره‌ای از دههٔ ۱۹۵۰ تا دههٔ ۱۹۸۰، نقش اصلی را در ریاضی کاربردی بازی کرده‌اند. هدف سه‌لایه است: حل مسائل عملی، فهمیدن ریاضیات نهفته در آنها و فراهم‌کردن بینش فیزیکی نسبت به این مسائل عملی. یک نمونهٔ موفق بسیار موجّه در این زمینه، مکانیک سیالات است. مکانیک سیالات نه‌تنها نیروی محرکهٔ اصلی در مطالعهٔ PDEها بوده است، بلکه این واقعیت که پژوهش در مکانیک سیالات تا حد زیادی به یک رشتهٔ محاسباتی تبدیل شده است نیز گواهی بر موفقیت این الگوریتم‌های عددی است که توسعه یافته‌اند. برای سال‌ها مطالعهٔ این PDEها و الگوریتم‌های عددی یکی از موضوع‌های اصلی ریاضی کاربردی بوده و امروزه همچنان یک حوزهٔ فعال است.

هنگامی که در دانشگاه کالیفرنیا (UCLA) دانشجوی تحصیلات تکمیلی بودم، با افتخار به ما گفته شد که ما در اردوگاه «ریاضی کاربردی به سبک کورانت» هستیم. این اصطلاح برای ایجاد تمایز با «ریاضی کاربردی به سبک بریتانیا» ابداع شده بود. هر دو اصطلاح روی مکانیک سیالات تمرکز داشتند. سبک بریتانیایی از بینش فیزیکی و مجانبی پیروی می‌کرد. رهبران این سبک، تِیلر، بَچلر، سی. سی. لین، لایتهیل و دیگران، نه‌تنها ریاضیدانان کاربردی بزرگی بودند، بلکه از نظریه‌پردازان دینامیک سیالات نیز

به‌شمار می‌آمدند. در نظر طرفداران این سبک، عموماً تحلیل دقیق و عددی از اهمیت بالایی برخوردار نبود. سبک کورانت از مطالعهٔ اعداد و قضیه‌ها، یا همان اثبات‌کننده‌های قضیه، پیروی می‌کرد. فلسفهٔ سبک کورانت این بود که مادامی که PDEهای زیربنایی و الگوریتم‌های عددی، قابل اعتماد هستند، از طریق محاسبات می‌توان چیزهای زیادی آموخت. با همهٔ اینها، فرایندهای فیزیکی بسیار پیچیده هستند و بدون محاسبه چیز زیادی دست‌گیرمان نمی‌شود. برخی از رهبران این سبک مانند فون‌نویمان، کورانت، فردریش و لَکس، نه‌تنها ریاضیدانان کاربردی بزرگی بودند، بلکه ریاضیدانان محض بزرگی نیز بودند. این واقعیت که اختلاف بین این دو مکتب، خصومت اصلی در ریاضی کاربردی محسوب می‌شد، از تسلط مکانیک سیالات در زمان آنها حکایت دارد.

آمار، پرچمدار جامعهٔ پژوهشی داده‌محور بوده است. به هر دلیلی، تا همین اواخر، آمار تا حدی به‌طور مستقل از ریاضی کاربردی و در واقع، به‌طور مستقل از ریاضیات، توسعه یافته است. آمار در گذشته بسیار به‌ندرت جزء برنامهٔ درسی گروه‌های ریاضی و یا ریاضی کاربردی بود و تنها در سال‌های اخیر است که فراخوان برای تغییر داده شده است. این بدان معنا نیست که جامعه ریاضی کاربردی به رویکرد داده‌محور علاقه نداشته است. به‌عکس، از اواخر دههٔ ۱۹۸۰، کار پژوهشی روی موجک و سنجش فشرده، پردازش سیگنال و تصویر یک بخش عمده از ریاضی کاربردی را به خود اختصاص داده است. در واقع، این نسخهٔ ریاضی کاربردی از رویکردی داده‌محور، از پربازده‌ترین حوزه‌های ریاضی کاربردی در سی سال گذشته بوده است. نه اینکه بگوییم مکانیک سیالات تنها حوزهٔ موفقیت‌آمیز برای ریاضیدانان کاربردی علاقه‌مند به PDEها بود. در واقع، برخی معتقد هستند که مکانیک جامدات نیز به همان اندازه موفق بوده است: با همهٔ اینها روش عناصر متناهی، یکی از مهم‌ترین موضوع‌های موفقیت‌آمیز در ریاضیات کاربردی، از مکانیک جامدات به‌وجود آمد. موضوع موفقیت‌آمیز دیگر، جبرخطیِ عددی است: کافی است ببینیم که محبوبیت نرم‌افزار متلب چقدر است تا تأثیر گستردهٔ آن را درک کنیم. نمونه‌های دیگری نیز از چنین داستان‌های موفقیت‌آمیزی وجود دارد.

۲. بحران در «ریاضی کاربردی به سبک کورانت»

متأسفانه برای نسل من از ریاضیدانان کاربردی به «سبک کورانت»، تسلط و موفقیت در مکانیک سیالات بیشتر یک چالش را ارائه کرد تا یک فرصت. کارهای اساسی برای PDEها و مکانیک سیالات، توسط نسل‌های قبل از ما انجام شده بود. ما در سر دوراهی قرار داشتیم که یا به مسائل باقیمانده مانند تلاطم بپردازیم و یا قلمروی جدید را تسخیر کنیم. هر دو نشان داده‌اند که دشوار هستند؛ جدای از بازتولید نوع موفقیتی که ریاضی کاربردی قبلاً در مکانیک سیالات داشته است. در واقع، پس از مکانیک سیالات، ریاضی کاربردی به سبک کورانت به بسیاری از رشته‌های علمی و مهندسی دیگر مانند علوم

مواد، شیمی، زیست‌شناسی، علوم اعصاب، علوم زمین و امور مالی گسترش یافته و موفقیت زیادی کسب کرده است. اما به‌طور کلی، میزان موفقیت در این حوزه‌ها با آنچه در مکانیک سیالات مشاهده کردیم، مطابقت ندارد. مشارکت‌های ما مورد استقبال قرار می‌گیرند اما به‌جای آنکه تحول‌آفرین باشند، بیشتر به‌صورت تدریجی انجام می‌شوند. در نتیجه برای پرداختن به موضوع‌های اساسی که با آن روبه‌رو هستند، دانشمندان یا پژوهشگران معمولاً مجبورند به تقریب‌های موردی متوسّل شوند که هم غیر قابل اعتماد و هم ناخوشایند هستند. این وضعیت در مکانیک کوانتوم، دینامیک مولکولی، دینامیک مولکولی درشت، مطالعات واکنش‌های شیمیایی، مدل‌های سیالات پیچیده، مدل‌های شکل‌پذیری، ساختار و دینامیک پروتئین، مدل‌سازی تلاطم، مسائل کنترل، برنامه‌ریزی پویا و غیره مشاهده می‌شود.

مشکل اصلی برای بیشتر، اگر نه همهٔ، این مسائل آن است که آنها ذاتاً مسائلی با ابعاد بالا هستند و نفرین مشقّت بُعدچندی[۱] آنها، همیشه گریبانمان را می‌گیرد. برای بیشتر مسائلی که در بالا بیان شد، زیاد بودنِ ابعاد نتیجهٔ ماهیت چندمقیاسه از مسئله است و تنها کورسوی امیدِ اندیشهٔ چندمقیاسه، مدل‌سازی چندفیزیکی بوده است. با دسته‌بندی درجات آزادی غیرضروری در مقیاس‌های کوچک، باید بتوان به‌طور مستقیم از مدل‌های ریزمقیاس قابل اعتمادتری استفاده کرد تا الگوریتم‌های بسیار کارآمدتری برای فرایندهای درشت‌مقیاس موردنظرمان ارائه داد. موفقیتِ مدل‌سازی چندمقیاسه هرچند بسیار امیدوار کننده است، تاکنون به دلایل زیر محدود شده است:

(۱) مدل‌های ریزمقیاس اغلب به آن اندازه قابل اعتماد نیستند. برای مثال، هنگام مطالعهٔ انتشار ترک، اغلب از دینامیک مولکولی به‌عنوان مدل ریزمقیاس استفاده می‌کنیم. اما دقت این مدل‌ها برای فرایندهایی که شامل شکستن پیوند است، اغلب زیر سؤال می‌رود.

(۲) اگرچه مدل‌سازی چندمقیاسه می‌تواند اندازهٔ شبیه‌سازی ریزمقیاس مورد نیاز را به‌طور قابل توجهی کاهش دهد اما این کار هنوز فراتر از توانایی فعلی ما است.

۳. یادگیری ماشین به کمک می‌آید

دلیل اصلی مشکلاتی که در بالا توصیف شد، توانایی محدود ما در مدیریت تابع‌هایی با متغیرهای زیاد است، و این دقیقاً همان جایی است که یادگیری ماشین می‌تواند متفاوت از ما عمل کند. با فراهم آوردن توانایی تقریب‌زدن تابع‌هایی با متغیرهای زیاد، مواردی که قبلاً غیرممکن تلقی می‌شدند، اکنون بسیار محتمل می‌شوند. پیش از یادگیری ماشین، حوزه‌ای که معمولاً از پسِ تابع‌هایی با متغیرهای زیاد بر می‌آمد، انتگرال‌گیری عددی بود. در فیزیک آماری، توانایی محاسبهٔ انتگرال تابع‌هایی با میلیون‌ها متغیر را تقریباً بدیهی می‌دانیم و فراموش می‌کنیم که این توانایی واقعاً قابل توجه است. این کار با الگوریتم

[۱] curse of dimensionality

مونت کارلو و فنون کاهش واریانس امکان‌پذیر است که در طی سالیان گذشته توسعه یافته‌اند. یک دلیل این است که برخلاف الگوریتم‌های مبتنی بر شبکه مانند قاعدهٔ سیمپسن، نرخ همگرایی الگوریتم مونت کارلو مستقل از بُعد است.

تقریب تابع‌ها در ابعاد بالا کاری بسیار دشوارتر است و موفقیت یادگیری ماشین در آن، به‌آسانی به‌دست نیامد [۶]. اگرچه مدل‌های شبکه‌های عصبی مدت‌ها پیش از آن کشف شده بودند، ظرفیت کامل آنها در تقریب تابع‌هایی با متغیرهای زیاد، اخیراً شناخته شده است. با وجود این، طی مدت زمان کوتاهی، شاهد دستاوردهای چشمگیر در حل چندین مسئله قدیمی با کمک یادگیری ماشین بوده‌ایم؛ امید می‌رود که چنین مسائلی نیز در آیندهٔ نزدیک حل شود.

ادغام یادگیری ماشین در ریاضی کاربردی، هر دو رشته را به روشی بنیادی تغییر می‌دهد. در ادامه، چند نمونهٔ خاص را توضیح خواهیم داد تا تأثیر این امر را در محاسبات علمی، مدل‌سازی، یادگیری ماشین و ریاضیات محض نشان دهیم.

۴. نظریهٔ کنترل و معادلات دیفرانسیل با مشتقات جزئی در ابعاد بالا

یکی از نخستین کاربردهای موفقیت‌آمیز یادگیری ماشین در محاسبات علمی، الگوریتمی مبتنی بر یادگیری عمیق برای مسائل کنترل در ابعاد بالا است [۵]. برای شروع بحث، جالب است بدانید که اصطلاح «مشقّت بُعدچندی» نخستین بار توسط ریچارد بلمن در برنامه‌ریزی پویا ابداع شد [۱]. در واقع، ابعاد معادلهٔ بلمن همان فضای حالتِ مسئله کنترل است: اگر به کنترل یک PDE علاقه‌مندیم، معادلهٔ بلمن بی‌نهایت بُعدی است. این امر کاربرد نظریهٔ کنترل «مبتنی بر اصل اول» را به‌شدت محدود کرده است و بسیاری از مسائل عملی باید با استفاده از گمانهٔ موردی حل شوند؛ درست مانند آنچه که برای مسائل چندپیکرهٔ کوانتومی انجام می‌شود.

در چارچوب کنترل طوقهٔ بسته، تابع خط‌مشی بهینه تابعی از حالت است. اگر این تابع خط‌مشی را با یک شبکهٔ عصبی پارامتری کنیم، شباهت بسیار خوبی بین کنترل تصادفی و یادگیری عمیق وجود دارد: تابع هزینه در مسئلهٔ کنترل، تابع زیان است؛ دستگاه دینامیکی برای مسئلهٔ کنترل نقش شبکهٔ ماندهٔ عمیق[۱] را بازی می‌کند؛ نوفه در دستگاه دینامیکی، نقش داده‌های آموزشی را بازی می‌کند که به ما امکان می‌دهد از الگوریتم گرادیان کاهشی تصادفی[۲] برای آموزش استفاده کنیم. با این الگوریتمِ مبتنی بر یادگیری عمیق، به‌طور معمول می‌توان از پسِ مسائل کنترل تصادفی در ابعاد صدها و حتی بالاتر برآمد [۵]. این روش به مسائل کنترل قطعی [۷] و PDEهای سهموی غیرخطی کلی نیز گسترش یافته است.

[۱] deep residual network [۲] stochastic gradient descent algorithm

این الگوریتم‌ها درها را برای مواجهه با مسائل کنترل در دنیای واقعی و PDEهای با ابعاد بالا گشوده‌اند. این یک امکان جدید مهیج است که باید بر اقتصاد، امور مالی، تحقیق در عملیات و یک دسته از رشته‌های دیگر تأثیر گذارد (و تا حدی نیز بر آنها تأثیر گذاشته است).

۵. مدل‌سازی به‌کمک یادگیری ماشین

در فیزیک ما به مدل‌های مبتنی بر اصل اول عادت کرده‌ایم. این مدل‌ها نه‌تنها کاربرد گسترده‌ای دارند، بلکه ساده و ظریف نیز هستند. معادلهٔ شرودینگر مثال خوبی در این زمینه است. متأسفانه همان‌طور که قبلاً نیز اشاره شد، حل مسائل عملی با استفاده از این مدل‌ها ممکن است کار بسیار دشواری باشد. بدین دلیل، جستجوی مدل‌های ساده یک موضوع دائمی در فیزیک و به‌طور کلی در علم بوده است. با وجود این، همان‌طور که در مدل‌های با تلاطم تجربه کرده‌ایم، اگر به تقریب‌های موردی متوسل نشویم، تهیهٔ چنین مدل‌های ساده‌ای اغلب بسیار دشوار خواهد بود.

یادگیری ماشین آماده است تا پیشرفتی عمده در توانایی ما برای توسعهٔ چنین مدل‌های مبتنی بر فیزیک ایجاد کند. این امر می‌تواند اتفاق افتد و در حال حاضر به سه روش مختلف اتفاق می‌افتد. نخست آنکه ابزار ناپیدایی را فراهم می‌کند که به‌کمک آن، می‌توان رؤیاهای مدل‌سازی چندمقیاسه را به واقعیت تبدیل کرد. دوم آنکه چارچوبی برای توسعهٔ مدل‌ها مستقیماً از روی داده‌ها فراهم می‌کند. سوم آنکه ابزاری بسیار قدرتمند برای تلفیق مدل‌های فیزیکی با مشاهده‌ها، چیزی شبیه داده‌گواری[1]، فراهم می‌کند.

با وجود این، برازش داده‌ها یک موضوع است، و ساخت مدل‌های فیزیکی قابل تعبیر و واقعاً قابل اعتماد، یک موضوع کاملاً متفاوت دیگر است. ابتدا در مورد موضوع تعبیرپذیری بحث می‌کنیم. به‌خوبی می‌دانیم که مدل‌های یادگیری ماشین به «جعبه‌های سیاه» مشهورند و این یک مانع روانی برای استفاده از یادگیری ماشین در کمک به توسعهٔ مدل‌های فیزیکی ایجاد کرده است. برای غلبه بر این مانع، توجه داشته باشید که تعبیرپذیری را باید به‌صورت نسبی نیز درک کرد. برای مثال، معادلهٔ اویلر برای دینامیک گازها را در نظر بگیرید. خودِ معادله‌ها به‌روشنی قابل تعبیر هستند، زیرا چیزی بجز پایستگی جرم، تکانه و انرژی را نشان نمی‌دهند. اما اینکه آیا جزئیات معادلهٔ حالت قابل تعبیر است یا نه، اهمیت کمتری دارد. در واقع، برای گازهای پیچیده، معادلهٔ حالت ممکن است دستورالعملی باشد که از درونیابی داده‌های تجربی با استفاده از اسپلاین‌ها به‌دست می‌آید. اینکه آیا ضرایب این اسپلاین‌ها قابل تعبیر هستند یا نه، برای ما اهمیتی ندارد. همین اصل باید برای مدل‌های مبتنی بر یادگیری ماشین نیز اعمال شود. در حالی که نقطهٔ شروع بنیادی این مدل‌ها باید قابل تعبیر باشد، درست مانند قوانین پایستگی در دینامیک گازها، صورت مفصل تابع‌های وارد شده به این مدل‌ها همگی قابل تعبیر نیستند. این تابع‌ها اغلب برخی روابط ساختاری را درست مانند معادلهٔ حالت برای دینامیک گازها نشان می‌دهند.

[1] data assimilation

اکنون به مسئله اعتمادپذیری می‌پردازیم. در شرایط آرمانی می‌خواهیم مدلی مبتنی بر یادگیری ماشین داشته باشیم که به‌اندازهٔ مدل‌های فیزیکی کلی مانند معادلهٔ ناویه–استوکس، برای همهٔ اهداف عملی قابل اعتماد باشد. برای تحقّق این امر، دو چیز مهم است. نخست اینکه مدل مبتنی بر یادگیری ماشین باید در همهٔ محدودیت‌های فیزیکی مانند موارد ناشی از تقارن‌ها و قوانین پایستگی صدق کند. دوم اینکه داده‌هایی که برای آموزش مدل استفاده می‌کنیم باید به‌اندازهٔ کافی پُربار باشند تا به‌درستی بتوانند تمام موقعیت‌های فیزیکی را که در عمل با آنها مواجه می‌شویم، نشان دهند. چون برچسب‌زنیِ داده‌ها تقریباً همیشه خیلی گران است، انتخاب یک مجموعه‌دادهٔ خوب که هم کوچک باشد و هم سرآمد، مؤلفه‌ای بسیار مهم برای توسعهٔ چنین مدل‌هایی است. در بخش بعد در این مورد بیشتر توضیح خواهیم داد.

این اندیشه‌ها قبلاً در مورد تعدادی از مسائل از قبیل دینامیک مولکولی و دینامیک گازهای رقیق با موفقیت استفاده شده است [۴]. در مورد دینامیک مولکولی، یادگیری ماشین همراه با محاسبات با کارایی بالا، امکان شبیه‌سازی دستگاه‌هایی با صدها میلیون اتم با دقت ابتدا به ساکن را فراهم کرده و نتیجهٔ آن پنج برابر کارایی بوده است (برای نمونه، [۴] را ببینید).

این تحوّلات جدید در حال حاضر کاملاً مهیج هستند. اما تأثیر مدل‌سازی به‌کمک یادگیری ماشین در حوزه‌هایی مانند زیست‌شناسی و اقتصاد که در آنها مدل‌سازیِ مبتنی بر اصل اول دشوار است، بیشتر ملموس خواهد بود. در حال حاضر برخی پیشرفت‌های مهیج در این زمینه‌ها در حال انجام است.

۶. مرزی جدید در یادگیری ماشین

ادغام یادگیری ماشین با ریاضی کاربردی همچنین منجر به ایجاد فرصت‌های جدید در یادگیری ماشین می‌شود. در اینجا دو مورد را بحث می‌کنیم.

۱.۶. یادگیری ماشین همزمان. در بیشتر چیدمان سنّتی یادگیری ماشین، داده‌های آموزشی یا از قبل تولید می‌شوند یا فقط مشاهده می‌شوند اما در مواردی که یادگیری ماشین برای حل مسائل محاسبات علمی یا علوم محاسباتی به‌کار گرفته می‌شود، معمولاً چنین نیست. در این شرایط، داده‌های آموزشی اغلب در حین اجرا تولید می‌شوند. برای ایجاد قیاسی در رابطه با مدل‌سازی چندمقیاسه که در آن، مدل‌سازی چندمقیاسهٔ متوالی با توجه به اینکه مدل‌های چندمقیاسه از قبل تولید شده باشند یا در حین اجرا، از مدل‌سازی چندمقیاسهٔ همزمان تشخیص داده می‌شود، این سبک یادگیری ماشین را *یادگیری ماشین همزمان* می‌نامیم. همان‌طور که قبلاً مشاهده شد، ایجاد یک مجموعه‌دادهٔ حداقل اما سرآمد، موضوعی کلیدی در یادگیری ماشین همزمان است.

در این راستا به شیوه‌ای کارآمد نیازمندیم تا در فضای حالت به کاوش بپردازیم و به معیاری برای تصمیم‌گیری نیاز است تا تعیین کند که آیا حالت جدید کشف‌شده، برچسب‌زنی شود یا خیر. یک نمونه در این زمینه الگوریتم EELT[۱] است که در [۴] پیشنهاد شده است.

۲.۶. صورت‌بندی خوش‌حالت از یادگیری ماشین. گذشته از قدرتمندیِ شگفت‌آور یادگیری ماشین مبتنی بر شبکهٔ عصبی، این روش بسیار شکننده است، زیرا کارایی آن با حساسیت به ابرپارامترهای موجود در مدل و الگوریتم آموزش بستگی دارد. در موارد بسیاری، تنظیم پارامترها هنوز هم کاملاً فنی است؛ هرچند با تجربهٔ اندوخته‌شده، این وضعیت به‌تدریج در حال بهبود است. این موضوع تا حدی به این دلیل است که در یادگیری ماشین، مدل‌ها و الگوریتم‌ها قبل از آنکه صورت‌بندی مسئله به‌طور دقیق مطالعه شود، ساخته می‌شوند. همین الان تصور کنید که اگر سعی کنیم بدون ساختن مدل‌های PDE از قبل، فرایندهای فیزیکی را مدل‌سازی کنیم، چه اتفاقی می‌افتد. در اصل، داشتن یک مدل PDE برای شروع کار و اطمینان از خوش‌حالت بودن آن یکی از مهم‌ترین درس‌هایی است که ما در ریاضی کاربردی به سبک کورانت آموخته‌ایم.

این امر به این پرسش می‌انجامد که آیا می‌توان به مدل‌های «مناسبِ» یادگیری ماشین دست یافت؟ انتظار می‌رود که اگر با صورت‌بندی‌های پیوستهٔ خوب شروع کنیم و سپس برای به‌دست آوردن مدل‌ها و الگوریتم‌های عملی آنها را گسسته کنیم، کارایی نسبت به انتخاب ابرپارامترها استوارتر خواهد بود. در این راستا برخی تلاش‌های آغازین در [۳] انجام شده است. جالب اینجاست که یک نتیجهٔ فرعی از کار انجام شده در [۳] نشان می‌دهد که مدل‌های شبکهٔ عصبی کاملاً طبیعی و اجتناب‌ناپذیر هستند، زیرا ساده‌ترین مدل‌های پیوسته و گسسته‌سازی‌ها همیشه به چیزی جز مدل‌های شبکهٔ عصبی یک شکل یا اشکال دیگر نمی‌انجامند. با وجود این، این روش پرداختن به یادگیری ماشین، مدل‌ها و الگوریتم‌هایی جدید ایجاد می‌کند. مهم‌تر از آن، این موضوع ما را تشویق می‌کند که به‌دنبال اصول اولیه باشیم و به ما امکان می‌دهد به خارج از کادر مدل‌های شبکهٔ عصبی نیز فکر کنیم.

قیاسی نزدیک در این زمینه، نوفه‌زدایی در پردازش تصویر است. روش استاندارد نوفه‌زدایی استفادهٔ مستقیم از پالایه‌های به‌دقت طراحی‌شده روی تصویر و دیدن رویداد است. این رویکرد، به‌ویژه با پالایه‌های پیشرفتهٔ مبتنی بر موجک، بسیار مؤثر بوده است. رویکرد دیگر نوشتن یک مدل ریاضی، معمولاً به‌صورت یک مسئله وردشی پیوسته، برای نوفه‌زدایی است. سپس مدل گسسته می‌شود و با الگوریتم‌های بهینه‌سازی مدل گسسته‌شده حل می‌شود. مدل‌های مشهور مامفورد–شاه[۲] و رودین–اوشر–فاطمی[۳] نمونه‌هایی از این نوع مدل‌های ریاضی هستند. اعتبار این مدل‌های ریاضی را می‌توان زیر سؤال برد اما روشن است که در آغاز، داشتن یک مدل ریاضی خوش‌تعریف مزیت خود را دارد. یک دلیل این است که چنین دیدگاهی به

[۱] Exploration-Examination-Labeling-Training [۲] Mumford-Shah [۳] Rudin-Osher-Fatemi

تبدیل پردازش تصویر به مسائل جالب PDE کمک کرده است. همچنین باعث ترغیب افراد به تفکّر دربارهٔ اصول بنیادی پردازش تصویر شده است؛ حتی اگر پیشرفت زیادی در این راستا حاصل نشده باشد.

امید این است که این شناخت و صورت‌بندی جدید ریاضی نه‌تنها به تقویت موفقیت فعلی یادگیری ماشین کمک کند، بلکه موفقیت آن را در طیف وسیعی از سایر رشته‌ها گسترش دهد. با همهٔ اینها، یادگیری ماشین در مورد تقریب تابع مسئله‌ای بسیار اساسی در ریاضیات است. داشتن راه‌های جدید برای نمایش و تقریب تابع‌ها که به‌ویژه در ابعاد بالا مؤثر هستند، مطمئناً باید تأثیری قابل‌توجه و گسترده داشته باشد.

۷. تحلیل ابعاد بالا

فقط حوزه‌های کاربردی نیستند که این تأثیر را تجربه می‌کنند، خودِ ریاضیات، به‌ویژه آنالیز ریاضی نیز تأثیر را احساس خواهد کرد. یادگیری ماشین انبوهی از مسائل جدید تحلیل ابعاد بالا را از تقریب تابع‌ها گرفته تا تقریب توزیع‌های احتمالی، دستگاه‌های دینامیکی و حل PDEها و معادله‌های بلمنی‌نما به‌وجود آورده است. مطالعهٔ این مسائل به‌طور حتم منجر به موضوعی جدید در ریاضیات به‌نام تحلیل ابعاد بالا می‌شود. در این راستا، انتگرال‌گیری در ابعاد بالا حوزه‌ای در ریاضیات است که مورد توجه جدی قرار گرفته است. تحلیل روش‌های مونت کارلو، به‌ویژه زنجیر مارکوف مونت کارلو، مدت زمانی طولانی یک حوزهٔ فعال در نظریهٔ احتمال و فیزیک ریاضی بوده است.

انتگرال‌گیری مقدماتی‌ترین مسئله در آنالیز ریاضی است. سؤال‌های پیشرفتهٔ خیلی بیشتری را می‌توان در مورد تابع‌ها، توزیع‌های احتمال، دستگاه‌های دینامیکی، حساب تغییرات و PDEها مطرح کرد. برای مثال، یک سؤال مهم توصیف پیچیدگی این اشیاء است. در یک سطح انتزاعی، پیچیدگی باید به‌وسیلهٔ دشواری‌ای تعریف شود که با آن شیء مفروض توسط اشیای مقدماتی ساده تقریب زده می‌شود. برای مثال، برای تابع‌ها، اشیای مقدماتی می‌تواند چندجمله‌ای‌ها، چندجمله‌ای‌های تکّه‌ای یا شبکه‌های عصبی باشند. برای توزیع‌های احتمال، اشیای مقدماتی می‌تواند ترکیبی از گاوسی‌ها باشند.

برای مثال، پیچیدگی تابع‌ها را درنظر بگیرید. به‌طور کلاسیک، این کار از طریق هموار انجام می‌شود؛ یعنی چند بار از تابع می‌توان مشتق گرفت. بسیاری از فضاهای تابعی سلسله‌مراتبی، مانند فضاهای C^k، فضاهای سوبولِف[۱] و فضاهای بِسُف[۲] به این صورت تعریف شده‌اند. در ابعاد پایین، این موضوع کاملاً منطقی است. در واقع می‌توان نشان داد که تابع‌ها در این فضاها، هنگامی که با برخی رده‌ها از تابع‌های مقدماتی مانند چندجمله‌ای‌های تکّه‌ای یا تابع‌های مثلثاتی تقریب زده می‌شوند، با نرخ همگرایی توصیف می‌شوند.

[۱] Sergei L. Sobolev [۲] Oleg V. Besov

این دست نتایج، دچار همان مشقّت بُعدچندی می‌شوند. در واقع، بیش از پیش معلوم شده است که مفاهیم مبتنی بر همواری راه درستی برای اندازه‌گیری پیچیدگی تابع‌ها در ابعاد بالا نیست. در عوض، باید پیچیدگی تابع‌ها در ابعاد بالا را به این صورت اندازه‌گیری کرد که آیا می‌توان آنها را با یک مدل شبکه‌نمای عصبی خاص به‌طور کارا تقریب زد. با این نگرش، فضای هیلبرت بازتولید هسته (RKHS)، فضای بَرون[۱]، فضاهای چندلایه و فضای ناشی از جریان[۲] به‌دست می‌آید که هر یک، به‌طور طبیعی با ردهٔ خاصی از مدل‌های یادگیری ماشین در ارتباط هستند.

در مورد PDEها در ابعاد بالا چه باید کرد؟ یک سؤال طبیعی این است که آیا می‌توانیم نظریه‌ای قاعده‌مند برای رده‌های معین PDEها در فضاهای تابعی مذکور توسعه دهیم. اگر می‌توانیم، این بدان معنی است که با استفاده از مدل یادگیری ماشین متناظر می‌توان جواب‌های این PDEها را به‌طور کارا تقریب زد. این مسئله برای معادلهٔ هَمیلتون–ژاکوبی–بلمن از اهمیت ویژه‌ای برخوردار است.

۸. ریاضی کاربردی به‌عنوان یک رشتهٔ علمی بالیده

آیا ریاضی کاربردی می‌تواند همانند ریاضی محض به موضوعی یکپارچه با تعدادی اندک از اجزای اصلی تبدیل شود؟ آیا می‌توان یک برنامهٔ درسی نسبتاً یکپارچه برای آموزش ریاضیدانان کاربردی ارائه کرد؟ پرداختن به این سؤال‌ها مدت‌ها دشوار بوده است. با نگاهی به گذشته، معلوم است که شرایط برای پاسخگویی فراهم نبوده است. از یک طرف، ریاضی کاربردی واقعاً بسیار متنوع است و تقریباً همهٔ رشته‌های علوم و مهندسی را تحت تأثیر قرار می‌دهد. جستجوی یگانگی و برنامهٔ درسی یکپارچه بدون شک کار دشواری است. از سوی دیگر، این واقعیت که اجزای اصلی مانند یادگیری ماشین در هستهٔ ریاضی کاربردی ناپیدا بوده‌اند به این معنی است که این جزء مهیا نبوده است. مثلاً تصور کنید ریاضی محض بدون جبر چگونه خواهد بود.

اوضاع تغییر کرده است. با ظهور یادگیری ماشین، همهٔ اجزای اصلی ریاضی کاربردی اکنون در جای خود قرار دارند. به عبارت دیگر، ریاضی کاربردی سرانجام آماده است تا به یک رشتهٔ علمی بالیده تبدیل شود. آری! مسیرهای جدید همچنان پدیدار می‌شوند اما دلایلی وجود دارد که معتقد باشیم مبانی کم‌وبیش در همان وضعیت باقی می‌مانند. این مبانی عبارتند از: مدل‌سازی (مبتنی بر اصل اول)، یادگیری و الگوریتم‌ها.

۱.۸. اجزای اصلی ریاضی کاربردی. جبر، آنالیز ریاضی، هندسه و توپولوژی اجزای اصلی ریاضی محض را تشکیل می‌دهند. مکانیک کلاسیک، مکانیک آماری، الکترومغناطیس و مکانیک کوانتومی نیز اجزای اصلی فیزیک هستند. اجزای اصلی ریاضی کاربردی کدامند؟ موارد زیر یک پیشنهاد است. این

[۱] Andrew R. Barron [۲] flow-induced space

به معنای سخن آخر در مورد این موضوع نیست، بلکه یک نقطهٔ شروع برای بحث بیشتر است. ریاضی کاربردی سه جزء اصلی دارد:

۱. مدل‌سازی مبتنی بر اصل اول که شامل مدل‌های (فیزیکی) و ابزارهای تحلیلی این مدل‌ها است. به بیان ساده، اولی در مورد فیزیک و دومی در مورد معادله‌های دیفرانسیل است. اصول نهفته در مدل‌های فیزیکی، اصول و قوانین بنیادی فیزیک از قبیل ساختار فیزیکی (مثلاً کلاسیک در مقابل کوانتومی، لَختی غالب در مقابل میرایی زیاد)، اصول وردشی، قوانین پایستگی و غیره هستند. این اصول اولیه برحسب مسائل وردشی یا معادله‌های دیفرانسیل صورت‌بندی شده‌اند. بنابراین برای مواجهه با این مسائل ریاضی به ابزارهای تحلیلی نیاز داریم. روش‌های مجانبی می‌توانند به‌سرعت ماهیت مسئله را به تصویر کشند و بینش مورد نیاز را به ما بدهند. قضیه‌های دقیق علاوه بر افشاندن نور به مسئله، می‌توانند به قرارگرفتن موضوع‌ها در یک پایهٔ محکم کمک کنند.

۲. روش‌های داده‌محور. با تفاوت بسیار مهم‌ترین بخش روش‌های داده‌محور یادگیری ماشین است که همچنین شامل آمار و پردازش داده‌ها (مثلاً تصویر) نیز هستند.

۳. الگوریتم‌ها. در این بحث الگوریتم‌ها هم برای کاربردهای مبتنی بر اصل اول و هم کاربردهای داده‌محور اهمیت دارند. خوشبختانه برای الگوریتم‌های هر دو حوزه وجه اشتراک زیادی وجود دارد. یک نمونه، الگوریتم‌های بهینه‌سازی است. این الگوریتم‌ها نه‌تنها در موفقیت یادگیری ماشین نقش محوری را ایفا کرده‌اند، بلکه خیلی از مدل‌های مبتنی بر اصل اول نیز به‌عنوان مسائل وردشی که برای آنها الگوریتم‌های بهینه‌سازی مورد نیاز هستند، صورت‌بندی شده‌اند.

۲.۸. برنامهٔ درسی و آموزش. بیشتر، اگر نه همهٔ، دانشگاه‌های برجسته دارای دوره‌های کارشناسی و تحصیلات تکمیلی ریاضیات محض نسبتاً بالیده هستند. تعدادی بسیار اندک از آنها برنامه‌های ریاضی کاربردی بالیده‌ای دارند. بدتر از آن، در بعضی موارد، دوره‌های ریاضی کاربردی به‌عنوان مجموعه‌ای از ترفندها و نه یک موضوع یکپارچه، تدریس می‌شود. یک مثال، درس «مدل‌سازی ریاضی» است. اگرچه این به معنای یک درس مقدماتی پایه برای ریاضی کاربردی است اما اغلب همانند مجموعه‌ای از مثال‌ها، بدون چشم‌اندازی منسجم، تدریس می‌شود. نبودِ برنامه‌های بالیدهٔ دورهٔ کارشناسی ریاضی کاربردی مهم‌ترین مانع برای ریاضی کاربردی است، زیرا توانایی ما در جذب استعدادهای جوان را به تأخیر می‌اندازد. با معلوم بودن اجزای اصلی ریاضی کاربردی، اکنون می‌توانیم یک برنامهٔ درسی یکپارچه برای ریاضی کاربردی طراحی کنیم. به‌طور طبیعی این برنامهٔ درسی بر پایهٔ سه مؤلفهٔ اصلی که قبلاً به آنها اشاره شد، قرار دارد. به‌طور خلاصه در مورد هر یک بحث می‌کنیم.

مدل‌سازی دو بخش دارد: اصول فیزیکی مدل‌ها و ابزارهای ریاضی برای تحلیل این مدل‌ها. مورد اول همانند مبانی فیزیک است که به ریاضیدانان آموزش داده می‌شود و مورد دوم تحلیل کاربردی شامل

ODEها و PDEها، حساب تغییرات، تحلیل توزیع‌های احتمال، مجانبی و آنالیز تصادفی است. هر درس را می‌توان با دوره‌ای یک‌ساله پوشش داد.

یادگیری به‌معنای تحلیل داده‌ها و شامل یادگیری ماشین، پردازش داده‌ها و آمار است. در حال حاضر درس‌هایی کامل برای پردازش داده‌ها و آمار وجود دارد که برای ریاضیدانان کاربردی مناسب است. وضعیت یادگیری ماشین متفاوت است و به‌طور معمول به سبکی مناسب برای متخصصین رایانه آموزش داده می‌شود. ما به روشی برای آموزش این موضوع به ریاضیدانان نیاز داریم. در این مرحله، دیدگاه ریاضی از یادگیری ماشین هنوز یک موضوع ناپخته است اما این وضعیت خیلی سریع در حال بهبود است. ما معتقدیم که به‌زودی یک دورهٔ مقدماتی ریاضیاتی مناسب برای درس یادگیری ماشین ایجاد می‌شود که می‌توان آن را در یک نیمسال تدریس کرد.

الگوریتم‌ها دو بخش دارند: الگوریتم‌ها برای اهداف پیوسته و الگوریتم‌ها برای اهداف گسسته. مورد اول با درس آنالیز عددی پوشش داده می‌شود که در گروه ریاضی ارائه می‌شود. مورد دوم با درس ریاضی الگوریتم‌ها یا گسسته پوشش داده می‌شود که معمولاً در گروه علوم کامپیوتر تدریس می‌شود. در یادگیری ماشین، این دو بخش در کنار هم قرار می‌گیرند، بنابراین مهم است که به سبکی یکپارچه‌تر به آنها آموزش داده شود. آماده‌سازی همهٔ این درس‌ها به تلاش بسیار زیادی نیاز دارد اما باید بتوانیم آن را عملی کنیم.

۹. ریاضی کاربردی به‌عنوان بنیاد پژوهش بین‌رشته‌ای

با اجرای چنین برنامه‌ای، ریاضی کاربردی بنیاد پژوهش بین‌رشته‌ای خواهد شد. با همهٔ اینها، مدل‌سازی، یادگیری و الگوریتم‌ها اجزای بنیادی تمام پژوهش‌های بین‌رشته‌ای نظری هستند. برنامهٔ ریاضی کاربردی که در بالا توصیف شد، به آموزش دانشجویان و همچنین سازماندهی برنامه‌های پژوهش بین‌رشته‌ای اسلوب خواهد داد و آن را نظام‌مند خواهد کرد. در صورت تحقق این امر نقطهٔ عطفی در تاریخ پژوهش بین‌رشته‌ای رقم می‌خورد.

همهٔ این برنامه‌ها زمان‌بَر است. یک دلیل این است که بایستی آموزش به دانشجویان جوان را از پایه شروع کنیم. اما پیش از آنکه فرصتی برای آموزش آنها داشته باشیم، باید بتوانیم آنها را مجذوب ریاضی کاربردی کنیم. چیزی که من به‌عنوان عضو هیئت علمی دانشگاه پرینستون با بیش از بیست سال سابقه، بسیار تحت تأثیر آن بوده‌ام این است که نظریهٔ اعداد چگونه توانسته افراد مستعد را به خود جذب کند. اکنون با کمال تعجب معتقدم که ریاضی کاربردی توانایی انجام همان کار را دارد. ریاضی کاربردی همهٔ ویژگی‌های اصلی که به‌ویژه برای دانشجویان جوان جذاب هستند را دارد: سادگی و ظرافت مسائل (مثلاً یادگیری ماشین) و چالشی که این مسائل مطرح می‌کنند (مثلاً تلاطم)، و همچنین این مزیت که یکی از

مسیرهای اصلی به سمت جذاب‌ترین تحوّلات جدید در علم و فناوری است. ما در آغاز مشاهدهٔ این تغییر در دوره‌های تحصیلات تکمیلی هستیم.

در تاریخ علم، دو دورهٔ زمانی وجود داشته که بیشترین تأثیر را در ریاضی کاربردی ایجاد کرده است. زمان نیوتن که در طی آن، مشخص شد ریاضیات باید زبان علم باشد، و زمان فون‌نویمان که در آن، پیشنهاد شد الگوریتم‌های عددی باید پل اصلی بین ریاضیات و علم باشند. اکنون زمان سوم در پیش است؛ زمانی که اجزای اصلی ریاضی کاربردی شامل مدل‌سازی، روش‌های داده‌محور و الگوریتم‌ها طوری در جای خود قرار می‌گیرند که ریاضی کاربردی را بنیادی برای پژوهش علمی بین‌رشته‌ای و نوآوری فناورانهٔ هیجان‌انگیز معرفی کنند و به‌راستی که زمانی هیجان‌انگیز است. بیایید همه با هم کار کنیم تا این را به واقعیت تبدیل کنیم!

مراجع

[1] Richard Bellman, *Dynamic programming*, Princeton University Press, Princeton, N. J., 1957.

[2] Paul A. Dirac, Quantum mechanics of many-electron systems, *Proc. Roy. Soc. London Ser. A*, **123** (1929), no. 792.

[3] Weinan E, Chao Ma, and Lei Wu, Machine learning from a continuous viewpoint, I., *Sci. China Math.,* **63** (2020), no. 11, 2233–2266.

[4] Weinan E, Jiequn Han, and Linfeng Zhang, Machine learning assisted modeling, *Physics Today*, **74** (2021), no. 7, 36–41 .

[5] Jiequn Han and Weinan E, *Deep learning approximation of stochastic control problems*, NIPS Workshop on Deep Reinforcement Learning, 2016.

[6] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, Deep learning, *Nature*, **521** (2015), 436–444.

[7] Tenavi Nakamura-Zimmerer, Qi Gong, and Wei Kang, *Adaptive Deep learning for high dimensional Hamilton-Jacobi-Bellman equations*, *SIAM J. Sci. Comp.*, **43** (2021), no. 2, 1221–1247.

شاهرخ اسماعیلی: سنندج، دانشگاه کردستان، گروه ریاضی

تارنما: http://sci.uok.ac.ir/esmaeili

رایانامه: sh.esmaeili@uok.ac.ir