مجلهٔ پژوهش فیزیک ایران

CC BY NC

## سرمقاله

# یادگیری ماشینی و علم فیزیک

داده‌ها به‌عنوان مجموعه‌ای از اطلاعات ارزشمند شناخته می‌شوند که با توسعهٔ زندگی صنعتی، تولید داده‌های مختلف نیز افزایش یافته‌است. یادگیری ماشینی، به‌عنوان یکی از زیرمجموعه‌های علم داده شناخته می‌شود که به توسعهٔ الگوریتم‌ها و روش‌هایی می‌پردازد که به کامپیوتر امکان می‌دهد بدون استفاده از دستورالعمل‌های صریح به یافتن طرح و همبستگی بین داده‌ها بپردازد.

این روش به ما امکان می‌دهد تا اطلاعات پنهان و الگوهای پیچیده‌ای را که در داده‌ها وجود دارد شناسایی کنیم.

در یادگیری ماشینی، کار برروی حجم عظیمی از داده‌ها انجام می‌گیرد. داده‌ها به چند بخش تقسیم می‌شوند. بخشی از داده‌ها توسط الگوریتم‌های موجود، آموزش داده‌می‌شود و نتایج آن با نتایج مشخص اولیهٔ آن بخش از داده‌ها مقایسه می‌شود. در این مرحله، آموزش الگوریتم انجام شده است. داده‌های آزمون چک می‌شوند. اگر خطایی در نتیجه مشاهده شود الگوریتم می‌تواند خود را تصحیح کند تا نتیجهٔ مدل با دقت خوبی مشابه نتایج داده‌های بخش آزمون شود. در این زمان، یادگیری ماشینی مشخص شده و برای استفاده از بقیهٔ داده‌های موجود برای پیش‌بینی و یافتن طرح‌ها و همبستگی‌های بین داده استفاده می‌شود. بنابراین نتایج هوشمند یادگیری ماشینی، اساساً مبتنی بر تطبیق الگو و کاوش داده‌ها به منظور دستیابی به نتیجه با کمترین دخالت انسانی است. برنامه‌های کاربردی یادگیری ماشین در زمینه‌های مختلفی مانند بینایی کامپیوتری، پردازش زبان، تشخیص الگو و پردازش تصویر توسعه یافته است. از موضوعات بسیار جذاب و مهم اخیر در یادگیری ماشین می‌توان به موارد زیر اشاره کرد:

یادگیری ماشینی چندمنظوره برای تلفیق الگوهای مختلف متشکل از داده‌هایی که ویژگی‌های متفاوتی دارند، شبیه‌سازی داده‌های ترکیبی و مصنوعی زمانی که در زمینهٔ خاصی داده و اطلاعات ورودی به اندازهٔ کافی در دسترس نباشد، و دوقلوهای دیجیتالی برای ساختن مدلی مجازی از یک دنیای واقعی با ویژگی‌های یکسان. دوقلوی دیجیتال کرهٔ زمین نمونهٔ بسیار جالب مورد آخر است که در این حالت مدل کرهٔ زمین با شرایط فیزیکی، آب و هوایی، و شرایط محیطی یکسان شبیه‌سازی می‌شود تا رویدادهای طبیعی قابل پیش‌بینی باشند.

علم داده، یادگیری ماشینی و علم فیزیک به طور جالبی با هم مرتبط هستند. در واقع، یکی از کاربردهای جالب یادگیری ماشین در فیزیک استفاده از مدل‌های پیش‌بینی برای تحلیل داده‌های تجربی و مشاهده شده است. این مدل‌ها که مبتنی بر روش‌های بدون ناظر[1] هستند، به ما کمک می‌کنند تا الگوهای پنهان در داده‌ها را شناسایی کرده و پیش‌بینی‌هایی در مورد رفتارهای فیزیکی سامانه داشته باشیم.

در فیزیک، روش‌های یادگیری ماشینی برای تجزیه و تحلیل داده‌های تجربی و مشاهده‌ای، شبیه‌سازی سامانه‌های

[1] Unsupervised

فیزیکی و پیش‌بینی بر اساس مدل‌های داده محور استفاده می‌شود. به‌عنوان مثال علم داده برای تجزیه و تحلیل داده‌های تجربی و شناسایی الگوها و روندها و استنتاج آماری آنها استفاده می‌شود. علم داده در اختر فیزیک برای تجزیه و تحلیل مجموعه داده‌های نجومی بزرگ مقیاس، مانند داده‌های به‌دست آمده از تلسکوپ‌ها و ماهواره‌ها برای مطالعهٔ اجرام و پدیده‌های آسمانی استفاده می‌شود. الگوریتم‌ها برای طبقه‌بندی و شناسایی اجرام ستارگان، کهکشان‌ها و سیاه‌چاله‌ها و تحلیل خواص آنها از قبیل جرم، دما و ترکیبات آنها استفاده می‌شود. در کیهان‌شناسی علم داده برای تجزیه و تحلیل داده‌های تابش پس‌زمینهٔ ریزموج پرتوهای کیهانی، بررسی‌های در مقیاس بزرگ از کهکشان‌ها و سایر مشاهدات تجربی به کار می‌رود که در مطالعهٔ ساختار، تکامل و ترکیب جهان و آزمایش مدل‌ها و نظریه‌های کیهان‌شناسی مهم هستند. علم داده در تجزیه و تحلیل داده‌های شتاب دهنده‌های ذرات مانند شتاب دهنده هادرون مورد استفاده قرار می‌گیرد. دانشمندان از الگوریتم‌های یادگیری ماشینی برای شناسایی و بازسازی ذرات از مقادیر وسیعی از داده‌های تولید شده توسط آشکارسازها استفاده می‌کنند که به درک جهان و برهم‌کنش‌های ذرات کمک می‌کند.

بنابراین علم داده و یادگیری ماشینی می‌تواند در علم فیزیک بسیار کمک کننده باشند. در این خصوص مثال‌های متعدد وجود دارد که خواننده را برای مرور موضوعات دیگر به مقالات زیر ارجاع می‌دهیم:

Nature Review Physics, 4, 761 (2022)
Nature, 620, 47 (2023).

حال پرسش اساسی و جالب این است که آیا علم فیزیک می‌تواند به یادگیری ماشینی و روش‌های آن کمک کند؟ جواب این سؤال مثبت است و علاوه بر این که محاسبات کوانتومی مبتنی بر کیوبیت کارایی بالایی برای کاربردهای یادگیری ماشین دارد، به چند مورد اخیر دیگر که فیزیک به یادگیری ماشین کمک کرده است اشاره می‌کنیم. اخیراً تراشه‌های نوری که در مقاله Nature Photonics, 17, 723 (2023) ارائه شده است، توسعه یافته‌اند که می‌توانند برای پردازش سریع‌تر الگوهای یادگیری ماشینی به‌کار گرفته شوند. تراشه‌های نوری سرعت پردازش چند برابر از ابرکامپیوترهای چت جی پی تی[1] دارند. انرژی مصرفی آن چند مرتبهٔ بزرگی کمتر از ابرکامپیوترهای مبتنی بر تراشه‌های الکترونی است. در این تراشه‌های نوری به جای این که از فرایند شبکهٔ عصبی مصنوعی مبتنی بر الکترون در فضای ۳ بعدی استفاده شود از پرتوهای نوری کمک گرفته می‌شود و در نتیجه در کاربرد شبکهٔ عصبی مصنوعی بسیار مفید است.

مثال دیگر در مورد چالش‌هایی است که یادگیری ماشینی با آن مواجه است. به عنوان مثال زمانی که داده‌ها از پیچیدگی خاصی برخوردارند و در عین حال تقارن‌هایی در آنها دیده می‌شود روش‌های معمولی یادگیری ماشین برای توجیه آنها با مشکل مواجه است. برای فائق آمدن بر آن نیاز است دانش بهتری از یادگیری ماشین داشته باشیم تا بتوان به نتایج مستخرج از آن اعتماد کرد.

روش جدید رگرسیون نمادین (به Symbolic Regression, Science Advances, 6, 16 (2020) رجوع شود) نوعی تحلیل رگرسیون است که جستجو در فضای توابع ریاضی انجام می‌گیرد تا مدلی را که از نظر دقت و هم از نظر سادگی با مجموعه‌ای از داده‌های معین مطابقت دارد ارائه دهد. مدل نهایی با کمینه کردن تابع اتلاف به‌دست می‌آید.

مثال دیگر شبکه‌های عصبی لاگرانژی است. در مدل‌های استاندارد، اعمال کردن تقارن‌ها در مدل‌های یادگیری ماشین ساده نیست. اما در شبکه‌های عصبی لاگرانژی می‌توان لاگرانژی‌های دلخواه را با استفاده از شبکه‌های عصبی پارامتری کرد. در این مدل نیازی به مختصات بنادی نیست و از این‌رو در شرایطی که اندازه حرکت بنادی نامشخص است یا به سختی محاسبه می‌شود روش شبکه‌های عصبی لاگرانژی عملکرد خوبی دارند.

در آخر باید اشاره کرد که آیندهٔ ارتباط تنگاتنگ بین یادگیری ماشینی و علم فیزیک روشن و پر امید است. با پیشرفت روزافزون فناوری و افزایش توانایی الگوریتم‌های یادگیری ماشینی، انتظار می‌رود که کاربردهای بیشتر و متنوع‌تری در فیزیک ظاهر شود. طراحی و توسعهٔ مواد جدید

[1] Chatgpt

با خواص چند منظوره، اندازه‌گیری‌های دقیق‌تر تجربی و حذف نوفه و خطا، کشف الگوهای جدید و ناشناخته پدیده‌های فیزیکی از نمونه مثال‌هایی است که انتظار می‌رود در آیندهٔ نزدیک به ثمر برسد.

رضا عسگری

سردبیر

۶ اردیبهشت ۱۴۰۲