

# معرفی کتاب: تئوری و الگوریتم‌های یادگیری ماشین

عباس محرابیان

دانشگاه بریتیش کلمبیا

AbbasMehrabian@gmail.com

۷ مرداد ۱۳۹۵

یادگیری ماشین<sup>۱</sup> در حال حاضر داغ‌ترین شاخه علوم کامپیوتر است و در سال‌های اخیر موفقیت‌های حیرت‌آوری کسب کرده است. یک نمونه جالب، برنامه کامپیوتری است که شرکت گوگل برای بازی GO طراحی کرده است. این برنامه داده‌های مربوط به ۱۰۰ هزار بازی واقعی را تحلیل کرده و چند ماه پیش موفق شد قهرمان جهان را شکست بدهد.<sup>۲</sup> (یک نکته جالب در مورد این برنامه اینست که می‌تواند با خودش بازی کند و از این کار چیز یاد بگیرد!) این بازی قوانین ساده‌ای دارد ولی فضای جستجوی آن بسیار وسیع است. تعداد حالت‌هایی که می‌تواند در آن رخ بدهد  $10^{170}$  است (در حالی که در شطرنج این عدد  $10^{40}$  است) و به همین دلیل کم‌تر کسی تصور می‌کرد کامپیوتر بتواند در این بازی بر انسان پیروز شود.

کلاس‌های درسی یادگیری ماشین شلوغ‌ترین کلاس‌ها در دانشگاه‌ها هستند و کلاس‌های رایگان اینترنتی متعددی هم در این زمینه به وجود آمده‌اند که یادگیری آن را برای هر دانشجوی علاقمندی که به اینترنت دسترسی دارد میسر می‌کند.<sup>۳</sup> بازار کار هم برای این رشته در آمریکای شمالی عالی است. در این مقاله ترجمه‌ای فارسی از مقدمه کتاب «تئوری و الگوریتم‌های یادگیری ماشین»<sup>۴</sup> را می‌آوریم. این کتاب مورد علاقه مترجم در این زمینه است و سه ویژگی کتاب که آن را منحصر به فرد می‌کنند عبارتند از:

۱. کتاب توسط دو ریاضی‌دان نوشته شده که از شاخص‌ترین پژوهش‌گران زمینه یادگیری ماشین نظری هستند و بنابراین از نظر ریاضی بسیار دقیق نوشته شده است.

۲. کتاب در سال ۱۳۹۳ چاپ شده و لذا بسیار جدید است.

---

<sup>۱</sup> machine learning

<sup>۲</sup> <https://deepmind.com/alpha-go>

<sup>۳</sup> برای دیدن فهرستی از این کلاس‌ها این پیوند را ببینید:

<https://www.class-central.com/subject/ai?sort=rating-up>

<sup>۴</sup> Understanding machine learning: from theory to algorithms

۳. نسخه الکترونیکی کتاب به صورت رایگان قابل دریافت است.<sup>۵</sup>

شایان ذکر است که تأکید این کتاب بر تحلیل ریاضی الگوریتم‌های یادگیری ماشین است، و بنابراین خیلی نظری است و ساختاری شبیه کتاب‌های ریاضی (تعریف، قضیه و اثبات) دارد اگرچه نگاهی هم بر کاربردهای یادگیری ماشین دارد. در نتیجه برای کسی که می‌خواهد آشنایی اولیه‌ای با یادگیری ماشین پیدا کند و یا کسی که جنبه‌های نظری و ریاضی برایش مهم نیست و می‌خواهد الگوریتم‌ها و کاربردهای یادگیری ماشین را یاد بگیرد شاید بهترین کتاب برای شروع نباشد.

**۱ مقدمه کتاب.** موضوع این کتاب یادگیری خودکار است، که اصطلاحاً به آن یادگیری ماشین گفته می‌شود. یعنی می‌خواهیم برنامه‌های کامپیوتری بنویسیم که از ورودی که به آن‌ها داده می‌شود چیز یاد می‌گیرند. معنای شهودی «یادگیری»، تبدیل تجربیات به تخصص است. ورودی یک الگوریتم یادگیری، داده‌های آموزشی<sup>۶</sup> است و خروجی اش یک برنامه کامپیوتری دیگر است که می‌تواند یک کار خاص را به خوبی انجام دهد. برای این که معنایی دقیق‌تری برای یادگیری ماشین پیدا کنیم، باید در مورد هر یک از پرسش‌های زیر بحث کنیم: داده‌های آموزشی که برنامه ما به آن دسترسی دارند چیستند و از کجا می‌آیند؟ اصولاً فرآیند یادگیری چگونه می‌تواند خودکار بشود؟ و در نهایت، چگونه می‌توانیم عملکرد الگوریتم یادگیری مان را ارزیابی کنیم؟

در این مقدمه ابتدا در این باره صحبت می‌کنیم که یادگیری اصولاً به چه معناست؟ سپس درباره کاربردهای یادگیری ماشین در عمل و ارتباطاتش با علوم دیگر صحبت می‌کنیم و در نهایت ساختار کتاب را شرح می‌دهیم.

**۲ یادگیری چیست؟** بحث را با دو مثال درباره یادگیری حیوانات آغاز می‌کنیم. وقتی یک موش به غذای جدیدی که قبلاً نخورده دست می‌یابد، ابتدا مقدار خیلی کمی از آن را می‌خورد، و صبر می‌کند تا ببیند دلش درد می‌گیرد یا نه. اگر دلش درد گرفت می‌فهمد غذا مسموم بوده و دیگر از آن غذا و غذاهایی با بو یا مزه مشابه نمی‌خورد. واضح است که این جا یک فرایند یادگیری در کار است: موش از تجربه‌ای که با غذا داشته درس می‌گیرد و درباره سلامت غذاهای مشابه یاد می‌گیرد. اگر در گذشته تجربه‌ای منفی با غذایی داشته، در آینده به آن دست نمی‌زند.

یک کاربرد ملموس و نابدیهی یادگیری ماشین، در تشخیص هرزنامه<sup>۷</sup> هاست. فرض کنید می‌خواهیم برنامه‌ای بنویسیم که یاد بگیرد تشخیص بدهد کدام نامه‌ها<sup>۸</sup> هرزنامه هستند و آن‌ها را دور بیندازد. اگر به مثال موش‌ها نگاه کنیم، راه حل ساده زیر به ذهن می‌رسد: نامه‌ها را مثل خوراکی‌ها ببینید، و هرزنامه‌ها را مثل خوراکی‌های مسموم. یک کاربر انسانی می‌تواند نقش سیستم هاضمه موش را بازی کند: هرگاه هرزنامه‌ای مشاهده شد، کاربر این موضوع را به برنامه اعلام می‌کند. سپس برنامه آن نامه را در جایی کپی می‌کند، و در آینده هرگاه نامه‌ای دریافت کرد که عیناً مشابه آن بود، آن را دور می‌ریزد و در صندوق پستی کاربر نمی‌آورد.

این ایده «یادگیری از طریق به خاطر سپردن» چندان قدرتمند نیست چرا که نمی‌تواند نامه‌های تازه را ارزیابی کند.

<sup>۵</sup> <http://goo.gl/Bvg4tM>

<sup>۶</sup> training data

<sup>۷</sup> spam

<sup>۸</sup> در این مقاله منظور از نامه، نامه الکترونیکی یا همان ای میل است.

ما از یک الگوریتم موفق انتظار داریم که با دیدن مثال‌های زیادی از هرزنامه‌ها و نامه‌های عادی به دانش کلی‌تری برسد و بتواند نامه‌های تازه و متفاوت را هم ارزیابی کند، کاری که به آن استنتاج استقرایی<sup>۹</sup> گفته می‌شود. مثلاً در مثال قبلی، موش اگر خوراکی جدیدی ببیند که بوی مزه «مشابه» غذاهای مسموم داشته باشد هم به آن دست نمی‌زند.

حال اگر بخواهیم الگوریتم قوی‌تری برای تشخیص و حذف هرزنامه‌ها طراحی کنیم، می‌توانیم مثلاً به نامه‌هایی که توسط کاربر هرزنامه تشخیص داده شده‌اند نگاه کنیم و ببینیم چه جور کلماتی در آن‌ها بیشتر تکرار شده‌اند، و آن کلمه‌ها را «ستاره‌دار» کنیم. در آینده هرگاه نامه‌ای دیدیم که حاوی تعداد زیادی کلمات ستاره‌دار بود آن را هرزنامه تشخیص می‌دهیم و دور می‌ریزیم. البته استنتاج استقرایی ممکن است به بی‌راهه هم برود. برای روشن شدن این موضوع مثال دیگری از یادگیری حیوانات، و این بار از پرندگان می‌آوریم. یک روانشناس تعدادی کبوتر گرسنه را داخل قفس بزرگی قرار داد، و سیستمی طراحی کرد که به صورت خودکار در بازه‌های زمانی مشخص غذا وارد قفس می‌کرد. وقتی غذا وارد می‌شد، هر کبوتر مشغول کاری بود، یکی به زمین نوک می‌زد، دیگری خودش را تمیز می‌کرد، سومی پرواز می‌کرد، و غیره. وقتی غذا می‌رسید کبوتر تصور می‌کرد آن کاری که در آن لحظه می‌کرده در رسیدن غذا مؤثر بوده، پس آن کار را بیشتر تکرار می‌کرد. این خودش باعث می‌شد دفعه بعدی که غذا می‌رسد، به احتمال بیشتری کبوتر مشغول آن کار خاص باشد، و این زنجیره باعث می‌شد کبوتر جداً رسیدن غذا را به انجام آن کار خاص ربط بدهد، و به صورت کورکورانه آن کار را تکرار کند.<sup>۱۰</sup> آیا می‌توان گفت کبوترها چیزی یاد گرفتند؟ پاسخ به وضوح منفی است. ما به صورت شهودی می‌توانیم بگوییم موش‌ها واقعاً یاد می‌گرفتند ولی کبوترها «توهم یادگیری» داشتند. ولی چه چیز یادگیری آن‌ها را از یادگیری موش‌ها متمایز می‌کند؟ این موضوع بخش بعدی است.

**۱.۲ دانش اولیه.** یکی از مسائل مهم در نظریه یادگیری ماشین اینست که معیارهایی ارائه کنیم و با کمک آن‌ها یادگیری واقعی را از یادگیری توهمی تمییز دهیم. برای تبیین یکی از این معیارها به آزمایش دیگری روی موش‌ها اشاره می‌کنیم. در این آزمایش وقتی به موش‌ها غذای خاصی می‌دادند به آن‌ها شوک الکتریکی وارد می‌کردند، و می‌خواستند ببینند که آیا این باعث می‌شود موش‌ها دیگر سراغ آن غذا نروند؟ با کمال تعجب دیده شد که هر چقدر این کار تکرار بشود، باز موش به سمت آن غذا می‌رود و از آن می‌خورد.<sup>۱۱</sup> گویی موش به صورت غریزی می‌داند که خوردن غذای بد ممکن است باعث مسمومیت بشود، ولی ربطی بین بد بودن غذا و شوک الکتریکی وجود ندارد!

نتیجه می‌گیریم تفاوت شاخصی که بین دو مثال موش و کبوتر وجود دارد، وجود یک جور «دانش اولیه» knowl- prior edge در موش‌هاست که فرآیند یادگیری آن‌ها را جهت‌دهی می‌کند: اگرچه کبوتر حاضر است هر گونه علتی را برای آمدن غذا بپذیرد، موش می‌داند که بد بودن غذا می‌تواند علت مسمومیت باشد ولی نمی‌تواند علت شوک الکتریکی باشد. بنابراین موش برخی هم‌زمانی‌های<sup>۱۲</sup> بین اتفاقات را مهم می‌داند و از آن‌ها یاد می‌گیرد و برخی دیگر را نادیده می‌گیرد.

inductive reasoning, inductive inference<sup>۹</sup>

<sup>۱۰</sup> برای اطلاعات بیشتر درباره این آزمایش به پیوند زیر مراجعه کنید  
<http://www.yorku.ca/pclassic/Skinner/Pigeon/>

<sup>۱۱</sup> برای اطلاعات بیشتر درباره این آزمایش به پیوند زیر مراجعه کنید:  
<http://goo.gl/PEvfbF>  
correlations<sup>۱۲</sup>

در حقیقت، می‌توان این موضوع را از نظر ریاضی ثابت کرد که هر الگوریتم یادگیری ماشینی، برای این که کار معناداری انجام دهد، لازم است که یک جور دانش اولیه در آن لحاظ شده باشد (که این کار توسط کاربر انسانی انجام می‌شود). دربارهٔ این موضوع در فصل پنجم کتاب مفصلاً بحث شده است. یکی از مرکزی‌ترین مباحث یادگیری ماشینی نظری، نظام‌بندی این دانش اولیه است و این که چطوری باید آن را در الگوریتم لحاظ کرد و از آن برای ارزیابی الگوریتم استفاده کرد. یک قاعده کلی اینست که هر چه این دانش اولیه—که در واقع همان مفروضات الگوریتم است—بیشتر باشد، یادگیری از تجربیات جدید ساده‌تر و سریع‌تر است. ولی در عین حال این باعث می‌شود الگوریتم انعطاف‌پذیری کم‌تری داشته باشد، چرا که این مفروضات خلاقیت الگوریتم را محدودتر می‌کنند.

**۳ یادگیری ماشین کجا به درد می‌خورد؟** در چه مواردی باید از یادگیری ماشین استفاده کنیم؟ دو دسته از مسائل هستند که برنامه‌های نوشته شده توسط انسان کارا نیستند و چاره‌ای نداریم جز این که وظیفهٔ تهیهٔ الگوریتم را به ماشین واگذاریم: یکی مسائل بسیار پیچیده، و دیگری مسائلی که ورودی‌شان مدام تغییر می‌کند.

یک سری از کارها هستند که ما انسان‌ها خیلی خوب به صورت ناخودآگاه انجام می‌دهیم، ولی اگر ازمان بپرسند نمی‌توانیم دقیقاً توضیح بدهیم که چطوری انجامشان می‌دهیم. مثلاً رانندگی، فهمیدن صحبت‌های یک شخص، و درک کردن تصاویر. تبدیل این مهارت‌ها به یک الگوریتم و نوشتن یک برنامهٔ کامپیوتری که بتواند یکی از این کارها را انجام بدهد خیلی سخت است. ولی امروزه الگوریتم‌های مبتنی بر یادگیری ماشین به وجود آمده‌اند که این کارها را به نحو احسن انجام می‌دهند.

دستهٔ دیگری از مسائل که روش‌های یادگیری ماشین در آن‌ها مؤثرند، تحلیل داده‌های بسیار بزرگ است: مثلاً پیش‌بینی وضعیت هوا با توجه به داده‌های هواشناسی و یا پیدا کردن یک صفحهٔ وب مناسب در پاسخ به درخواست یک کاربر (کاری که یک موتور جستجو انجام می‌دهد). داده‌های دیجیتال در دنیا مدام بیشتر و بیشتر تولید می‌شوند و تحلیل آن‌ها و استخراج دانش از آن‌ها کاری بسیار مهم و در عین حال غیرممکن برای یک انسان است. افزایش غول‌آسای قدرت پردازش کامپیوترها باعث شده الگوریتم‌های یادگیری ماشین بتوانند در این زمینه بسیار موفق عمل کنند.

فرض کنید می‌خواهیم برنامه‌ای بنویسیم که کاربر بتواند با او حرف بزند و برنامه حرف او را بفهمد. واضح است که یک الگوریتم ثابت نمی‌تواند این کار را انجام بدهد، چرا که هر کاربر لهجه، تن صدا و لحن خود را دارد. الگوریتم‌های یادگیری ماشین، که اصولاً بر اساس کسب تجربه و انطباق با محیط کار می‌کنند، در چنین مواردی راهگشا هستند. یک مثال دیگر تشخیص هرزنامه‌هاست که قبلاً به آن اشاره کردیم. پنج سال پیش محتوای هرزنامه‌ها با امروز خیلی فرق داشت؛ الگوریتم یادگیری ماشین می‌تواند در طول زمان بر اساس تغییر ورودی‌هایش، با شرایط جدید تطبیق پیدا کند، بنابراین می‌تواند در امر تشخیص هرزنامه‌ها موفق‌تر از یک الگوریتم ثابت باشد.

**۴ زیرشاخه‌های یادگیری ماشین.** «یادگیری» مفهومی بسیار گسترده است و یادگیری ماشین هم به زیرشاخه‌های متعددی تقسیم می‌شود. در حقیقت می‌توان آن را از زاویه‌های مختلفی طبقه‌بندی کرد. در این جا چند طبقه‌بندی از یادگیری ماشین ارائه می‌دهیم، با این هدف که چهارچوب این کتاب را روشن‌تر کنیم.

**۱.۴ یادگیری با نظارت در مقابل یادگیری بدون نظارت.** در مثال تشخیص هرزنامه، فرضمان بر اینست که کاربری وجود دارد که می‌تواند تشخیص بدهد کدام نامه‌ها هرزند و کدام‌ها نیستند، و الگوریتم یادگیری می‌تواند از این اطلاعات استفاده کند. به یادگیری این چینی، یادگیری بانظارت<sup>۱۳</sup> گفته می‌شود.

مثال‌های دیگری از مسائل هم هستند که در آن‌ها ناظری وجود ندارد. واضح‌ترین مثال مسئله خوشه‌بندی<sup>۱۴</sup> حجم وسیعی از داده‌ها به گروه‌هایی است که اعضای هر گروه شبیه به هم باشند. به این یادگیری، یادگیری بدون نظارت<sup>۱۵</sup> گفته می‌شود، چرا که ناظری وجود ندارد که تشخیص بدهد خروجی الگوریتم درست بوده یا نه. به همین دلیل معمولاً ارزیابی عملکرد الگوریتم‌های بدون نظارت دشوارتر است.

یک حالت بینابینی هم وجود دارد که مثال خوبش رانندگی است. در این مورد خروجی الگوریتم مجموعه‌ای از دستورالعمل‌های کوچک (مثل چرخاندن فرمان به راست، فشار دادن ترمز و غیره) است که هر کدام تک‌تک قابل ارزشیابی توسط ناظر نیستند؛ در این جا تنها خروجی نهایی الگوریتم (که یا اتومبیل را سالم به مقصد می‌رساند یا نه) است که قابل ارزیابی است. دقت کنید که این کار الگوریتم را بسیار سخت می‌کند چرا که اگر تصادف کند تشخیص این که کدام یک از اعمالش اشتباه بوده کار ساده‌ای نیست. به این مدل یادگیری، یادگیری تقویتی<sup>۱۶</sup> گفته می‌شود.

**۲.۴ یادگیری فعال در مقابل یادگیری غیرفعال.** در یادگیری با نظارت، یادگیرنده می‌تواند به صورت فعال با ناظر ارتباط داشته باشد و داده‌های آموزشی تولید کند و به ناظر بدهد، یا این که صرفاً به مثال‌هایی که از ناظر به او می‌رسد قانع باشد. در مثال تشخیص هرزنامه، یادگیرنده فعال<sup>۱۷</sup> خودش نامه‌هایی تولید می‌کند و از کاربر می‌خواهد هرز بودن آن‌ها را تعیین کند، ولی یادگیرنده غیرفعال<sup>۱۸</sup> چنین امکانی ندارد. واضح است که یادگیرنده غیرفعال کار سخت‌تری پیش رو دارد.

**۳.۴ نقش محیط در یادگیری.** وقتی در مورد یادگیری آدمیزاد حرف می‌زنیم، مثلاً یادگیری دانش‌آموزی در سر کلاس، همیشه معلم دلسوزی را در نظر می‌آوریم که به انسان در یادگیری کمک می‌کند و مثال‌های آموزنده برایش می‌زند. وقتی زیست‌شناسی پدیده‌های طبیعی را بررسی می‌کند، محیط نقش معلم دلسوز را بازی نمی‌کند، بلکه در حقیقت طبیعت مثال‌های آموزشی را به صورت کاملاً تصادفی تولید می‌کند. یادگیری از مثال‌های تصادفی شاخه‌ای از یادگیری ماشین است که یادگیری آماری<sup>۱۹</sup> نام دارد. در سر دیگر طیف می‌توان حالتی را متصور شد که محیط سعی می‌کند مثال‌های آموزشی را طوری تولید کند که یادگیرنده را گمراه کند. مثلاً در مثال تشخیص هرزنامه، می‌توان تصور کرد کسی که هرزنامه را تولید می‌کند سعی می‌کند نامه‌اش طوری باشد که الگوریتم آن را یک نامه عادی تشخیص بدهد.

<sup>۱۳</sup> supervised learning

<sup>۱۴</sup> clustering

<sup>۱۵</sup> unsupervised learning

<sup>۱۶</sup> reinforcement learning

<sup>۱۷</sup> active learner

<sup>۱۸</sup> passive learner

<sup>۱۹</sup> statistical learning

**۴.۴ یادگیری برخط در مقابل یادگیری برون خط.** یادگیری برخط<sup>۲۰</sup> به این معناست که ورود داده‌های جدید و خروجی‌های الگوریتم در طول زمان در هم تنیده هستند. واضح‌ترین مثال آن الگوریتمی است که در بازار بورس، سهام خرید و فروش می‌کند. این الگوریتم باید در هر لحظه تصمیم به خرید یا فروش سهام بگیرد و این طور نیست که اول مقدار زیادی داده دریافت کند و بتواند همه تصمیماتش را پس از تحلیل همه آن داده‌ها بگیرد. الگوریتم بازار بورس در آغاز کارش داده‌های اندکی در اختیار دارد و لذا کارش سخت‌تر است.

ما در این کتاب همه انواع یادگیری ماشین را در بررسی نمی‌کنیم. عمده بحث کتاب در حیطه یادگیری آماری غیرفعال با نظارت و به صورت برون خط<sup>۲۱</sup> می‌باشد. به عنوان یک مثال بارز در این حیطه، فرض کنید می‌خواهیم الگوریتمی طراحی کنیم که وجود یک بیماری خاص را در افراد به اطلاعات پزشکی آنان ربط دهد. داده‌هایمان عبارتند از اطلاعات پزشکی مربوط به تعداد زیادی آدم تصادفی که مستقلاً جمع‌آوری شده‌اند، و برای هر آدم این را هم می‌دانیم که آن بیماری را داشته یا خیر. البته در یکی دو فصل کتاب درباره یادگیری برخط و یادگیری بدون نظارت (به ویژه مسئله خوشه‌بندی) هم بحث می‌کنیم.

**۵ ارتباط با دیگر شاخه‌های علم.** یادگیری ماشین شاخه‌ای بین‌رشته‌ای است و به ویژه با شاخه‌های ریاضی مثل آمار، نظریه اطلاعات، نظریه بازی‌ها و بهینه‌سازی ارتباط گسترده‌ای دارد. از آن جا که هدف آن در نهایت برنامه‌نویسی کامپیوترهاست، طبیعی است که آن را زیرشاخه‌ای از علوم کامپیوتر بدانیم. به یک معنا یادگیری ماشین زیرشاخه‌ای از هوش مصنوعی<sup>۲۲</sup> است، چون به هر حال تبدیل تجربه به دانش را می‌توان نوعی هوشمندی محسوب کرد. ولی باید در نظر داشت که در هوش مصنوعی به معنای کلاسیک، هدف صرفاً شبیه‌سازی هوش انسانی است، در حالی که هدف یادگیری ماشین استفاده از قدرت کامپیوتر برای تکمیل هوش انسانی است. در برخی از موارد از این هم فراتر می‌رویم و کارهایی با یادگیری ماشین انجام می‌شود که بسیار فراتر از توانایی انسان است. یک مثال بارز توانایی کار با پایگاه داده‌های بسیار بزرگ و تشخیص الگو در آن‌هاست که مثلاً منجر به موتور جستجوی قدرتمندی مثل گوگل شده است.

فرایند یادگیری معمولاً از طریق کار با داده‌هایی که به صورت تصادفی تولید شده‌اند انجام می‌شود. الگوریتم یادگیری این داده‌ها را تحلیل می‌کند و سعی می‌کند الگوهایی در آن‌ها پیدا کند، و در مورد توزیع احتمالاتی آن‌ها اطلاعاتی کسب کند. از همین جا معلوم می‌شود یادگیری ماشین با آمار ریاضی ارتباط نزدیکی دارد، ولی این دو رشته تفاوت‌های جدی هم دارند، که در این جا مختصراً و در فصل پنجم مفصلاً به آن‌ها اشاره می‌کنیم.

مثلاً فرض کنید یک متخصص قلب پس از مطالعه بیماران بسیاری، حدس می‌زند که شاید ارتباطی بین سیگار کشیدن و احتمال بیماری قلبی وجود داشته باشد. آن‌گاه این دکتر داده‌های خود را به یک آمارپژوه می‌دهد که کارش تأیید یا رد این حدس است. در مقابل، هدف یک الگوریتم یادگیری ماشین اینست که داده‌های مربوط به بیماران را بدون هیچ پیش‌فرضی تحلیل کند و فهرستی از دلایل احتمالاً مرتبط با بیماری قلبی ارائه کند. بنابراین امید اینست که این الگوریتم الگوهایی در داده‌ها پیدا کند و بر اساس آن‌ها حدس‌هایی بزند که کاربر انسانی از آن‌ها غافل مانده است.

online learning<sup>۲۰</sup>  
offline, batch<sup>۲۱</sup>  
artificial intelligence<sup>۲۲</sup>

تفاوت دیگر با آمار اینست که در یادگیری ماشین ما با الگوریتم‌های کامپیوتری سروکار داریم و طبیعتاً سرعت عملکرد این الگوریتم‌ها برایمان اهمیت زیادی دارد. این نکته وقتی با داده‌های غول‌آسا مثل صفحات وب کار می‌کنیم پررنگ‌تر هم می‌شود. تفاوت سوم اینست که در آمار معمولاً رفتار حدی پارامترها (وقتی اندازه داده به بی‌نهایت میل می‌کند) است که بیشترین اهمیت را دارد، حال آن‌که در یادگیری ماشین باید دقت الگوریتم را در حالتی که اندازه داده متناهی است هم از نظر ریاضی تحلیل کنیم.

**۶ ساختار این کتاب.** بخش اول کتاب (فصل‌های ۲ تا ۸) اصول ریاضی یادگیری ماشین را ارائه می‌دهند. بقیه کتاب بر اساس این بنیان‌های ریاضی بنا شده است. می‌توان بر اساس این بخش درس کوتاهی در زمینه پایه‌های نظری یادگیری ماشین ارائه کرد.

در بخش دوم کتاب (فصل‌های ۹ تا ۲۰) مهم‌ترین و پرکاربردترین الگوریتم‌های یادگیری ماشین با نظارت را معرفی می‌کنیم. می‌توان از این مطالب این بخش برای ارائه یک درس «آشنایی با یادگیری ماشین» برای دانشجویانی که لزوماً به پایه‌های ریاضی علاقمند نیستند استفاده کرد.

در بخش سوم کتاب (فصل‌های ۲۱ تا ۲۵) از محدوده یادگیری با نظارت فراتر می‌رویم و مقداری درباره مدل‌های دیگر یادگیری (مثل یادگیری بر خط و بدون نظارت) و جنبه‌های دیگر یادگیری ماشین بحث می‌کنیم.

بخش چهارم کتاب (فصل‌های ۲۶ تا ۳۱) برای کسانی است که قصد دارند در زمینه یادگیری ماشین نظری پژوهش جدی انجام دهند. در این بخش ریاضیات پیشرفته‌تری را معرفی می‌کنیم که این پژوهش‌گران به آن نیاز دارند. نهایتاً در پیوست کتاب ابزارهای ریاضی مورد نیاز (به ویژه از نظریه احتمال و جبر خطی) برای مطالعه کتاب را مرور می‌کنیم.