



سازمان‌ها پیش از آن‌که تصمیم بگیرند یادگیری ماشینی را در عمل مورد استفاده قرار دهند باید با کم و کیف این فناوری آشنا شوند تا ناخواسته قربانی یک پیاده‌سازی اشتباه نشوند.

اشتباهات

یادگیری ماشینی به داستان‌هایی که در فیلم‌های علمی‌تخیلی به آن‌ها اشاره می‌شود، محدود نمی‌شود. این فناوری همچون منبع انرژی پیرامون هر یک از فناوری‌هایی که در زندگی روزمره ما مورد استفاده قرار می‌گیرند، قرار دارد. تشخیص صدا از سوی سیری یا آلكسا، تشخیص خودکار برچسب‌ها در فیسبوک، توصیه‌هایی که از سوی آمازون و اسپاتی‌فای ارائه می‌شوند همگی بر مبنای یادگیری ماشینی قرار دارند. این حضور به اندازه‌ای موفقیت‌آمیز بوده است که بسیاری از سازمان‌ها علاقه‌مند شده‌اند تا از الگوریتم‌های یادگیری ماشینی برای افزایش بهره‌وری شبکه خود استفاده کنند. در حقیقت تعدادی از این سازمان‌ها به منظور افزایش بهره‌وری سامانه‌های شناسایی و بهینه‌سازی گسترده شبکه‌های خود از مدت‌ها قبل از یادگیری ماشینی استفاده کرده‌اند. اما هر فناوری منجمله یادگیری ماشینی می‌تواند همانند یک شمشیر دو لبه عمل کرده و اگر به شکل نادرستی پیکربندی شود، باعث ویرانی یک شبکه شود. به همین دلیل شرکت‌ها قبل از پذیرش این فناوری باید با راه‌هایی که ممکن است یادگیری ماشینی زمینه‌ساز سقوط آن‌ها شود آشنا شوند و پیش از آن‌که مجبور شوند برای جبران خسارت به عقب گام برداشته و عملیات احیا را اجرا کنند با کم و کاستی‌های این فناوری آشنا شوند. رومان سینایو، مهندس هوشمندی امنیتی نرم‌افزار در ژوپیتر نتورکس راهکارهایی را معرفی کرده است که از بروز اشتباهاتی که به واسطه یادگیری ماشینی ممکن است یک سازمان را تهدید کند، ممانعت به عمل می‌آورد.

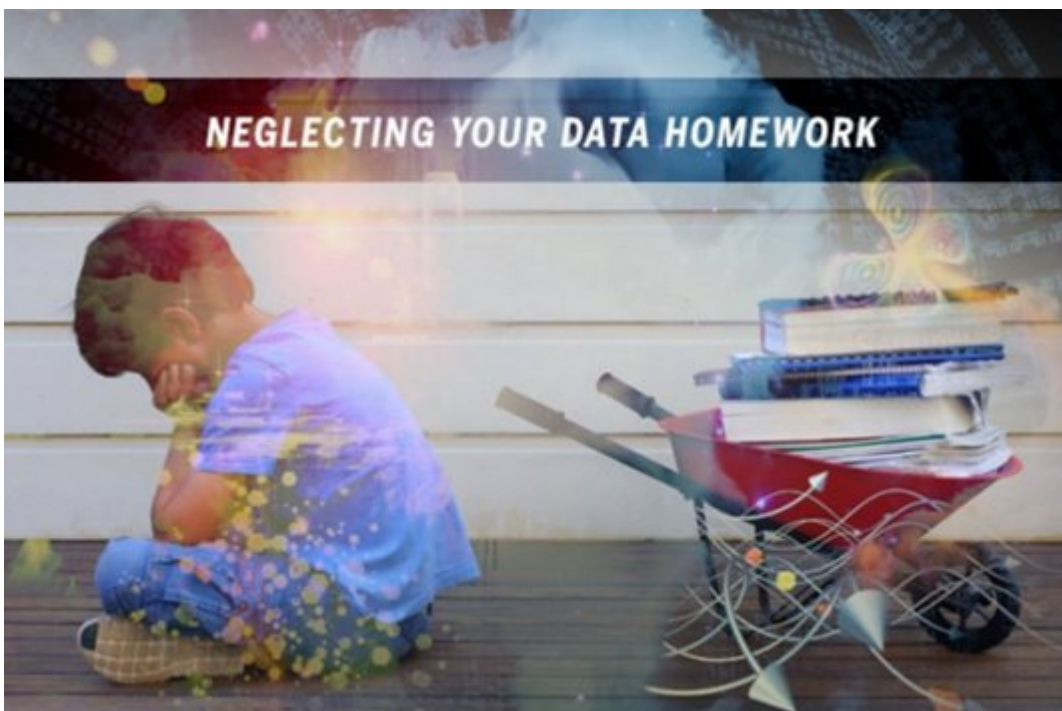
به رفتارهای ناپایدار غیرمنتظره دقت کنید



شگفت‌انگیز است، موضوعی که یک کامپیوتر فکر می‌کند مهم است و به آن واکنش نشان می‌دهد، زمانی که توسط عامل انسانی مورد بررسی قرار می‌گیرد، یک موضوع بی اهمیت تشخیص داده می‌شود. به همین دلیل، ضروری است به این نکته توجه داشته باشیم که بسیاری از متغیرهای زمینه و نتایج بالقوه ممکن است بعد از استقرار یادگیری ماشینی خود را نشان دهند و در عمل به یک تهدید ناخواسته تبدیل شوند.

اجازه دهید این موضوع را با یک مثال روشن کنیم. یک مدل آموزش دیده است تا تصاویر مربوط به وسایل نقلیه سبک و کامیون‌ها را در دو گروه مجزا از هم طبقه‌بندی می‌کند. اما اگر تمامی تصاویر مربوط به کامیون‌ها در شب گرفته شده باشد و تمامی تصاویر مربوط به ماشین‌ها در روز گرفته شده باشد، این احتمال وجود دارد که مدل این‌گونه تشخیص دهد که هر تصویر مربوط به یک ماشین که در شب گرفته شده است ممکن است یک کامیون باشد. آدرس‌دهی درست متغیرهای کلیدی و نتایجی که در دوره‌های آزمایشی کسب می‌شوند، کمک خواهند کرد تا حد امکان از بروز رفتارهای ناخواسته و غیرمنتظره ممانعت به عمل آوریم و راه‌حل مناسبی برای آن‌ها ارائه کنیم.

غفلت از مشق شب، عدم درک درست داده‌ها



به منظور ساخت یک مدل آموزش دیده، ابتدا باید یک درک اولیه به دست آید و در ادامه داده‌هایی که در فرآیند تحلیل مورد استفاده قرار می‌گیرند، جمع‌آوری شوند. این اطلاعات برای تعیین متغیرها و نتایج بالقوه‌ای که عملکرد یک الگوریتم را تحت تاثیر خود قرار می‌دهند، ضروری هستند. علاوه بر این، اگر یک مدل اصل طبقه‌بندی داده‌ها را فراموش

کرده باشد، این امکان وجود دارد که با بهترین داده‌ها که قادر به ارائه یک راه‌حل ایده‌آل هستند آموزش نیندند.

توسعه، آزمایش و در نهایت اجرای کردن مدل



برای آن‌که بتوانید مدلی را تولید کنید که کاربردی و مفید باشد، باید در ابتدا یک ساختار داده‌ای آموزش‌دهنده با کیفیت در اختیار داشته باشید. پیش از آن‌که یک الگوریتم یادگیری ماشینی داخل یک سازمان استقرار یابد، علم داده‌ها پیشنهاد می‌کند که ابتدا مدل را با مجموعه‌ای از داده‌ها آزمایش کنید تا از عملکرد قابل قبول آن اطمینان حاصل کنید. داده‌ها پیش از آن‌که در قالب یک

ساختار داده‌ای که بتواند قابلیت خودیادگیری را در اختیار یک مدل قرار دهد، آماده شوند باید دو فرآیند را پشت سر بگذارند. اول آن‌که با تلاش بسیار مجازی‌سازی شده باشند و دوم آن‌که تحت نظارت قرار گرفته باشند. علم داده‌ها ممکن است یک مدل را در سریع‌ترین زمان ممکن مورد آزمایش قرار دهد، اما برای این منظور ممکن است از مجموعه داده‌هایی استفاده کند که شاید در دنیای واقعی الگوریتم یادگیری ماشینی هیچ‌گاه با آن‌ها روبرو نشود. برای این منظور ضروری است برای متغیرهای انتخاب شده داده‌های کافی در اختیار داشته باشید تا فرآیند آزمایش الگوریتم به درستی انجام شود. تغذیه یک مدل با اطلاعات بیشتر در این مرحله به بهبود عملکرد کمک فراوانی کرده و تضمین می‌کند یک مدل یادگیری ماشینی در عمل باعث افزایش بهره‌وری محیط تولید شده و عملکرد واحدهای عملیاتی را بهبود می‌بخشد.

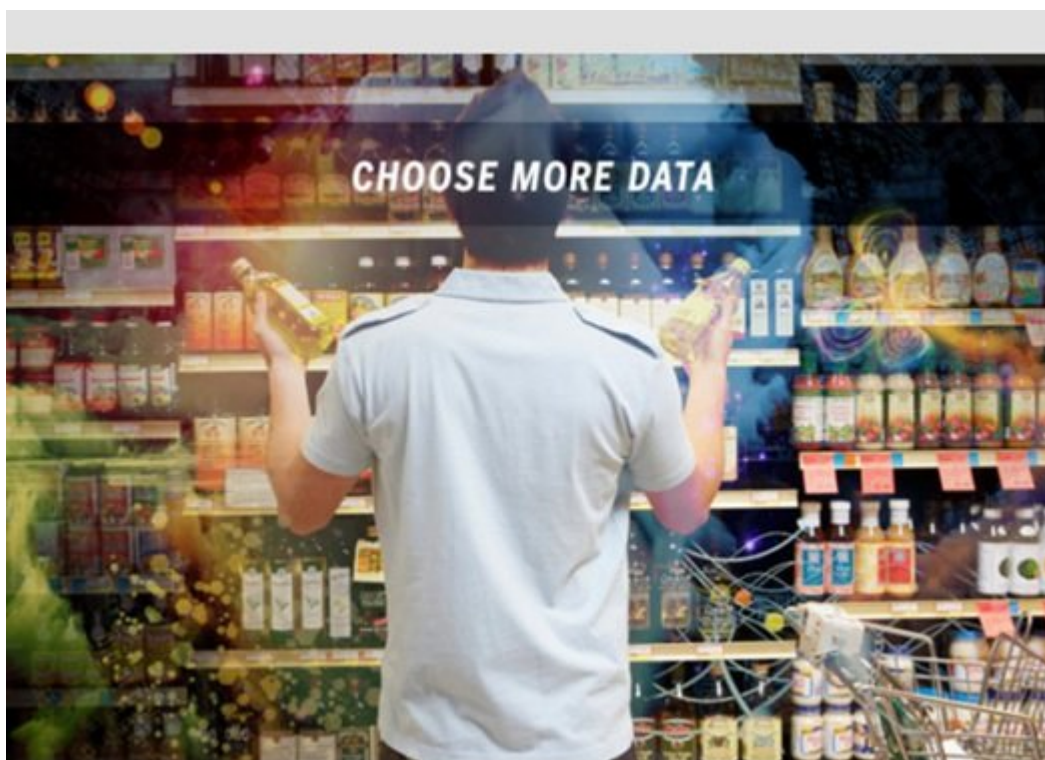
نادیده گرفتن اشتباهات



ممکن است
پیش از آن که
یک پروژه به
هدف نهایی
خود نزدیک
شود با موانع
جدیدی روبرو
شود که باعث
به وجود آمدن
اشتباهات
بالقوه‌ای شود.
در یک نمونه
مشهور، یک
شرکت بزرگ
یک روبات
رسانه‌ای را
طراحی کرد.
این روبات برای
تقلید الگوهای
زبان و تکامل

بهرتر قابلیت‌های تعاملی طراحی شد. اما در عمل کاربران با یک روبات بحث‌برانگیزی روبرو شدند که صحبت‌های جنجال‌برانگیزی را به زبان می‌آورد. به طوری که شرکت در نهایت مجبور شد بخشی از طراحی روبات را که در ارتباط با یادگیری رفتارها بود مجدداً بازطراحی کند. اما در نهایت مجبور شد بعد از گذشت 24 ساعت به کار این روبات برای همیشه پایان دهد. هر پروژه یادگیری ماشینی را نمی‌توان به صورت عمومی عرضه کرد یا نمی‌توان به کاربران اجازه داد تا به صورت آزاد و بدون کنترل به هر پروژه یادگیری ماشینی دسترسی پیدا کنند و داده‌ها را دستکاری کنند. اما آگاهی از محیطی که الگوریتم در آن وارد می‌شود باعث می‌شود تا از اشتباهات بالقوه جلوگیری شود.

داده‌های بیشتری انتخاب کنید



زمانی که یک مدل را به لحاظ عملکرد مورد آزمایش قرار می‌دهید، ممکن است نتایجی که مدنظر دارید را دریافت نکنید. برای حل این مشکل دو راهکار پیش روی شما قرار دارد. الگوریتم یادگیری ماشینی را بهتر و دقیق‌تر طراحی کنید یا داده‌ها بیشتری جمع‌آوری کنید. اضافه کردن داده‌های بیشتر به مهندسان کمک می‌کند تا درباره محدودیت‌های عملکردی الگوریتم درک بهتری به دست آورند. اگر بتوانید داده‌های

زیادی را جمع‌آوری کنید، آن‌گاه الگوریتم شما به شکل کافی تغذیه می‌شود و در نتیجه قادر خواهید بود نتایج درستی را به دست آورید. این‌کار به شما کمک می‌کند تا مجبور نباشید الگوریتم خود را از نو مورد بازطراحی مجدد قرار دهید.

قاعده خارج از اصول طراحی نکنید



نوع ویژه‌ای از الگوریتم یادگیری ماشینی که چند وقتی است مورد توجه قرار گرفته و البته کاربردهای عملی آن نیز با موفقیت به اثبات رسیده است، مدل یادگیری تجمعی است. فرآیندی که چند مدل هوش محاسباتی برای حل یک مشکل با یکدیگر ترکیب می‌شوند، یک

نمونه از مدل یادگیری تجمعی رگرسیون لجستیکی است. این روش‌های یادگیری تجمعی می‌تواند به بهبود عملکرد پیش‌بینی‌ها در مقایسه با روش‌های مشابه منجر شوند.

تاریخ انتشار:
16 آذر 1395

نشانی منبع: <https://www.shabakeh-mag.com/artificial-intelligence/5702>