یادگیری ماشین در امنیت سایبری

تیر ۱۴۰۱

ترجمه شده توسط شرکت مهندسی امن ارتباط سینداد

رویارویی تهدیدهای سایبری با یادگیری ماشین

امروزه استفاده مهاجمان از روش‌های پیشرفته و نیز به کارگیری راهکارهای منسوخ‌شده امنیت سایبری در سازمان ها، منجر به ایجاد آسیب‌پذیری‌ها و حفره‌های امنیتی متعدد شده‌اند و درنتیجه کسب‌وکارها پیوسته به دنبال راه‌حل‌های پویا و مؤثر هستند.

طبق یک نظرسنجی، ۹۳ درصد از مدیران اذعان داشته‌اند که حاضرند ۲۲ درصد بیش از حالت عادی برای سیستم هایی با امنیت بالا هزینه پرداخت می‌کنند.

در نیمه اول سال ۲۰۲۰، مهاجمان حدود ۳۶ میلیارد مدرک سازمانی را سرقت وا افشا کردند.

هزینه‌های پاک‌سازی باج افزارها در سیستم‌های سازمانی و بازگردانی رویه کسب‌وکار به حالت قبل از حمله سایبری و دیگر سایر هزینه‌ها از ۷۶۱,۰۰۰ دلار در سال ۲۰۲۰، به ۱.۸۵ میلیون دلار در سال ۲۰۲۱ افزایش پیداکرده است.

بر اساس تحقیقات انجام‌شده، فقط ۵٪ از اسناد سازمان‌ها به‌درستی در برابر نفوذهای سایبری محافظت دارند.

جرائم سایبری همگام با تکنولوژی پیشرفت می‌کنند و در مقابله با تهدیدها و حمله‌های سایبری ضرورت دارد همیشه یک‌قدم از آن‌ها جلوتر باشیم. فناوری‌های متعددی وجود دارند که در برابر حملات سایبری ایستادگی و محافظت می‌کنند؛ یکی از فناوریهای مورداستفاده زیاد و به‌سرعت در حال رشد، یادگیری ماشین است. قابلیت یادگیری ماشین در شناسایی فعالیت‌های کاربران و سیستم‌ها می‌باشد و به‌واسطه‌ی آن رفتارهای غیرعادی و حمله‌های احتمالی تشخیص داده می‌شوند.

فناوری یادگیری ماشین نیازمند وجود داده‌های تولیدشده پیشین است تا بتواند الگوهای رایج در آن‌ها را استخراج نماید. بااین‌حال، یادگیری ماشین یک راه حل قطعی در مقابله با تمامی حمله‌های سایبری نیست، زیرا جرائم سایبری به‌طور پیوسته در حال تغییر و گسترش هستند.

یادگیری روش‌ها و تکنیک‌های حمله، در مقابله با تهدید های سایبری و رخداد احتمالی آنان در آینده،الزامی است. در سالیان اخیر، یادگیری ماشین تبدیل به بخش حیاتی از امنیت سایبری شده است؛ به‌ویژه در انجام وظایف امنیتی روزمره‌ای مانند طبقه‌بندی، الگو یابی و پیش‌بینی، که اقداماتی بنیادین در شناسایی و پاسخ‌دهی به حملات سایبری هستند.

یادگیری ماشین چیست؟

یادگیری ماشین زیرمجموعه‌ای از هوش مصنوعی به شمار می‌آید که از الگوریتم‌های تکامل‌یافته تشکیل‌شده است. یادگیری ماشین از سوابق داده‌ها الگو گیری می‌کند و در پردازش داده‌های فعلی استفاده می‌کند و به پیشرفت الگوریتم‌ها سرعت می‌بخشد.

عملکرد یادگیری ماشین از طریق مطالعه و شناسایی داده‌های مرتبط و مناسب ارتقاء پیدا می‌کند و به میزان داده‌های باکیفیت بیشتر، خروجی به‌دست‌آمده از آن‌ها نیز بهتر می‌شود و سیستم‌هایی که داده‌های تولیدی زیادی به همراه خوددارند، پتانسیل بالاتری برای پیشرفت یادگیری ماشین فراهم می‌نمایند.

۳ روش اصلی یادگیری ماشین

۱. یادگیری نظارت‌شده

این‌یک روش غیرمتداول از یادگیری ماشین است و معمولاً در طبقه‌بندی و رگرسیون داده‌ها و ویژگی‌هایشان استفاده می‌شود. در این روش، الگوریتم‌ها بر مبنای داده‌های قدیمی تولید می‌شوند و بدین ترتیب، با توجه به نتایج به‌دست‌آمده، سیستم هدف حمله فرآیند یادگیری خودکار را شروع می‌کند. به این روش از یادگیری ماشین، یادگیری نظارت‌شده می‌گوییم.

حوزه‌های فعال امنیت سایبری از این نوع یادگیری ماشین، در شناسایی حمله‌های شناخته‌شده، استفاده می‌کنند و بیشترین کاربرد آن هنگامی است که راهکارهای قدیمی امنیت در تشخیص تهدیدها ناتوان هستند. بنابراین، به‌منظور تأثیرگذاری این روش، باید داده‌های مناسب و باکیفیت به آن داده شود.

۲. یادگیری نظارت‌نشده

یادگیری نظارت‌نشده بسان چگونگی یادگیری نوزادان است. آن‌ها رخدادهای در وقوع اطراف خود را تماشا می‌کنند و تلاش دارند الگویی معنادار از آن بسازند .

این روش یادگیری ماشین، معمولاً در دسته‌بندی و کاهش حجم ابعاد داده‌ها استفاده می‌شود. در روش یادگیری بدون نظارت، الگوریتم مقدار زیادی از داده‌های بدون شناسه را می‌خواند تا سطح میزان خطاها را کاهش دهد و داده‌های جمع‌آوری‌شده با معنادار شدن به استخراج الگوی اصلی اطلاعات کمک کند.

این نوع از یادگیری ماشین که در امنیت سایبری به کار می‌رود، برای تشخیص ناهنجاری‌ها و تجزیه‌وتحلیل رفتارهایی استفاده می‌شود که شاخصه خاصی برای شناسایی ندارند.

برای مثال اگر یک کاربر در یکی از واحدهای سازمان که تابه‌حال از دستورات Shell استفاده نکرده است، ناگهان شروع به استفاده از آن کند، به‌عنوان رفتاری غیرمعمول و ناهنجار توسط یادگیری ماشین شناسایی می‌شود.

۳. یادگیری تقویتی

یادگیری تقویتی، زیرمجموعه‌ای از هوش مصنوعی است که در آن از تکنیک آزمون‌وخطا یا اکتشاف استفاده می‌شود.رویکرد یادگیری تقویتی(Reinforcement learning)همانند روش‌هایی است که انسان برای انجام فعالیت‌های جدید به کار می‌برد. این روش از یادگیری ماشین در پس‌زمینه بسیاری از ابزارهای کاربردی و فناوری‌های برجسته به کار می‌رود که به‌عنوان نمونه، می‌توان به اتومبیل‌های خودران اشاره کرد.

هدف یادگیری تقویتی بهبود مستمر سیستم است؛ به این صورت که از شکست‌ها درس گرفته و با اعمال نتایج مثبت و منفی مرتبط با عملیاتی که انجام می‌دهد، هر بار پیشرفت می‌کند.

امروزه روش یادگیری تقویتی در امنیت سایبری راهکار بسیار قدرتمندی است. برای مثال، یادگیری تقویتی برای تست نفوذ به‌APIهای خاص استفاده می‌شود و همچنین برای تأمین امنیت به‌صورت خودکار در نرم‌افزارهای امنیتی تحت شبکه نیز استفاده می‌شود.

موارد رایج استفاده از ماشین یادگیری در امنیت سایبری

یادگیری ماشین پیوسته در حال تکامل است و موارد استفاده از آن نیز روزبه‌روز در حال افزایش است.

موارد ارزشمند در به‌کارگیری یادگیری ماشین عبارت‌اند از:

۱-شناسایی و توقف باج افزارها و دیگر فعالیت‌های مخرب

باج افزار یک نوع سرقت سایبری است که در آن از تکنیک‌های رمزنگاری پیشرفته استفاده می‌شود و با رمزنگاری داده‌های سیستم مهاجم طلب باج جهت بازیابی آن‌ها میکند.

درصورتی‌که یک حمله‌ی باج افزاری موفقیت‌آمیز باشد، کاربران کنترلشان بر روی سیستم‌ها را از دست می‌دهد و طی یک درخواست از سوی باج افزار متحمل پرداخت هزینه هنگفت به گرداننده‌ی آن باج افزار می‌شوند. برای

پیشگیری از وقوع این رخداد، از یادگیری ماشین برای شناسایی الگوهای رفتاری باج افزارها استفاده می‌شود. یادگیری ماشین، شیوه‌های مهندسی‌شده پیشرفته‌ای را به کار می‌گیرد تا رفتار انواع باج افزارها را شناسایی کند و بتواند با تهدیدهای باج افزاری روز مقابله کند.

۲-مقابله با تهدیدهای ناشناس

سیستم‌های امنیتی لازم است توانایی مقابله با تهدیدهای پیشرفته و ناشناخته مانند باج افزارهای مبتنی بر آسیب‌پذیری‌های روز صفر(zero-day) را داشته باشند. با پیشرفت باج افزارها، راه محافظت در برابر آن‌ها نیز باید تغییر پیدا کند.

با توسعه و گسترش باج افزارهای امروزی و رشد بی‌وقفه آن‌ها، سازمان‌ها به قابلیت‌های پویا و اثربخشی مانند یادگیری ماشین نیاز دارند تا هم‌زمان با پیشرفت جرائم سایبری، به آمادگی لازم جهت محافظت در برابر آن‌ها دست پیدا کنند.

۳-استخراج اطلاعات مفید داده‌های انبوه

متخصصان امنیتی از مهم‌ترین منابعی هستند که وظیفه شناسایی حمله‌های مخرب، تجزیه‌وتحلیل ترافیک شبکه، بررسی لاگ‌ها و رخدادهای امنیتی نقاط پایانی و تشخیص آسیب‌پذیری‌ها را بر عهده‌دارند. یکی از مهم‌ترین مشکلات امنیتی سازمان‌ها، ایجاد هشدارهای امنیتی بیش‌ازحد است که درصد بالایی از این هشدارها، کاذب (False Positive) هستند. وقتی تحلیلگران امنیتی با حجم زیادی از هشدارهای امنیتی کاذب روبه‌رو میشوند، ممکن است هشدارهای ضروری را نادیده بگیرند. در چنین شرایطی، یادگیری ماشین باقابلیت شناسایی داده‌های خطا و کاهش چشم‌گیر آن‌ها، راه‌حل مؤثری برای افزایش توانایی تجزیه‌وتحلیل هر سازمانی است.