مجله یافته های علوم تربیتی و آموزشی

کاربرد هوش مصنوعی در تشخیص و پیشگیری از حملات سایبری

طیبه پرهیزکاری ، پردیس اخوت پور

کارشناسی ارشد مدیریت فن آوری اطلاعات، دانشگاه آزاد اسلامی

Tararp_81@Yahoo.com

کارشناسی نرم افزار کامپیوتر، دانشگاه آزاد اسلامی

Pardisokhovat@gmail.com

MDOI-02-2026.0371
MNR-385-2-D8EBDF

چکیده

گسترش روزافزون فناوری اطلاعات، رایانش ابری، اینترنت اشیا، سامانه‌های صنعتی و خدمات دیجیتال موجب شده است که وابستگی سازمان‌ها و جوامع به زیرساخت‌های سایبری به شکل قابل توجهی افزایش یابد. هم‌زمان با این تحول، تهدیدات سایبری نیز از حملات ساده و مبتنی بر روش‌های شناخته‌شده به حملاتی پیچیده، چندمرحله‌ای، خودکار و سازگار با محیط تبدیل شده‌اند. مهاجمان سایبری با استفاده از بدافزارهای پیشرفته، حملات فیشینگ، باج‌افزارها، حملات منع خدمت توزیع‌شده، سرقت اعتبارنامه‌ها، سوءاستفاده از آسیب پذیری های ناشناخته و روش های مهندسی اجتماعی، تلاش می‌کنند محرمانگی، یکپارچگی و دسترس‌پذیری اطلاعات و خدمات را مختل کنند. در چنین شرایطی، روش‌های سنتی امنیت سایبری که عمدتاً بر امضاهای شناخته‌شده، قواعد از پیش تعیین‌شده و تحلیل دستی متکی هستند، به‌تنهایی پاسخ‌گوی حجم و پیچیدگی تهدیدات جدید نیستند.

هوش مصنوعی و به‌ویژه یادگیری ماشین و یادگیری عمیق، ظرفیت قابل توجهی برای ارتقای توانایی سازمان‌ها در شناسایی، تحلیل، پیش‌بینی و مقابله با تهدیدات سایبری ایجاد کرده‌اند. سامانه‌های مبتنی بر هوش مصنوعی می‌توانند حجم گسترده‌ای از داده‌های شبکه، گزارش‌های ثبت رویداد، رفتار کاربران، فرایندهای سیستم، فایل‌ها و شاخص‌های تهدید را تحلیل کرده و الگوهای غیرعادی را در زمان کوتاه شناسایی کنند. این فناوری همچنین در تشخیص بدافزار، شناسایی فیشینگ، تحلیل رفتار شبکه، کشف نفوذ،

پیش‌بینی نقاط آسیب‌پذیر، اولویت‌بندی هشدارها، تحلیل تهدیدات و خودکارسازی واکنش به رخدادها کاربرد دارد.

با وجود مزایای قابل توجه، استفاده از هوش مصنوعی در امنیت سایبری با چالش‌هایی مانند کیفیت و کمبود داده، هشدارهای کاذب، قابلیت توضیح‌پذیری، حملات خصمانه علیه مدل‌های یادگیری ماشین، مسموم‌سازی داده‌های آموزشی، حفظ حریم خصوصی و احتمال سوءاستفاده مهاجمان از خود فناوری هوش مصنوعی همراه است. ازاین‌رو، هوش مصنوعی نباید جایگزینی کامل برای متخصصان امنیت تلقی شود، بلکه باید به‌عنوان بخشی از یک معماری دفاعی چندلایه مورد استفاده قرار گیرد. این مقاله با رویکردی توصیفی ـ تحلیلی، کاربردهای اصلی هوش مصنوعی در تشخیص و پیشگیری از حملات سایبری، مزایا و محدودیت‌های آن و الزامات پیاده‌سازی مؤثر این فناوری را بررسی می‌کند.

کلیدواژه‌ها: هوش مصنوعی، امنیت سایبری، حمله سایبری، یادگیری ماشین، یادگیری عمیق، تشخیص نفوذ، پیشگیری از حملات، تشخیص بدافزار، فیشینگ، تحلیل رفتار.

۱. مقدمه

تحول دیجیتال در سال‌های اخیر ساختار بسیاری از سازمان‌ها، کسب‌وکارها و خدمات عمومی را دگرگون کرده است. امروزه اطلاعات مالی، ارتباطات سازمانی، اطلاعات مشتریان، سامانه‌های مدیریتی و حتی برخی زیرساخت‌های حیاتی از طریق شبکه‌های رایانه‌ای و سامانه‌های متصل به اینترنت مدیریت می‌شوند. در نتیجه، هرگونه اختلال یا نفوذ به این سامانه‌ها می‌تواند پیامدهای اقتصادی، اجتماعی و حتی امنیتی قابل توجهی به همراه داشته باشد.

امنیت سایبری در ساده‌ترین تعریف، مجموعه‌ای از فرایندها، فناوری‌ها و اقدامات مدیریتی برای حفاظت از سامانه‌ها، شبکه‌ها، برنامه‌ها و داده‌ها در برابر تهدیدات دیجیتال است. یکی از اجزای اساسی امنیت سایبری، تشخیص نفوذ است. مؤسسه ملی استاندارد و فناوری ایالات متحده آمریکا ، تشخیص نفوذ را فرایند پایش رویدادهای رخداده در یک سامانه یا شبکه و تحلیل آن‌ها برای یافتن نشانه‌های احتمالی رخدادهای امنیتی تعریف می‌کند.

با افزایش حجم داده‌های تولیدشده در شبکه‌ها، تعداد رخدادهای امنیتی و سرعت تغییر تاکتیک‌های مهاجمان، تحلیل دستی این داده‌ها دشوار شده است. یک مرکز عملیات امنیتی ممکن است روزانه هزاران یا حتی میلیون‌ها رویداد را دریافت کند، در حالی که تنها بخش محدودی از آن‌ها واقعاً نشان‌دهنده یک حمله سایبری هستند. بنابراین یکی از مهم‌ترین مسائل امنیت سایبری مدرن، تشخیص سریع رفتارهای مشکوک در میان حجم عظیم داده‌های عادی است.

هوش مصنوعی می‌تواند در این زمینه نقش مهمی ایفا کند. الگوریتم‌های یادگیری ماشین قادرند از داده‌های گذشته الگوهای رفتاری را یاد بگیرند و در مواجهه با داده‌های جدید، احتمال وقوع رفتار مخرب را ارزیابی کنند. یادگیری عمیق نیز امکان تحلیل داده‌های پیچیده‌تر و کشف روابطی را فراهم می‌کند که ممکن است با قواعد ساده قابل شناسایی نباشند.

از این منظر، رابطه هوش مصنوعی و امنیت سایبری دوطرفه است: از یک سو هوش مصنوعی ابزار قدرتمندی برای دفاع سایبری محسوب می‌شود و از سوی دیگر باید خود هوش مصنوعی در برابر حملات محافظت شود.

۲. مفهوم هوش مصنوعی و جایگاه آن در امنیت سایبری

هوش مصنوعی به مجموعه‌ای از روش‌ها و فناوری‌ها اطلاق می‌شود که به سامانه‌های رایانه‌ای امکان انجام وظایفی را می‌دهند که معمولاً به نوعی از توانایی‌های شناختی انسان نیاز دارند؛ از جمله تشخیص الگو، تصمیم‌گیری، پیش‌بینی، طبقه‌بندی و تحلیل داده.

در حوزه امنیت سایبری، مهم‌ترین شاخه‌های هوش مصنوعی عبارت‌اند از:

یادگیری ماشین؛

یادگیری عمیق؛

پردازش زبان طبیعی؛

تحلیل رفتار؛

شبکه‌های عصبی مصنوعی؛

روش‌های تشخیص ناهنجاری؛

سیستم‌های خبره؛

مدل‌های مولد و زبانی.

یادگیری ماشین یکی از مهم‌ترین فناوری‌های مورد استفاده در امنیت سایبری است. در این روش، الگوریتم به جای آنکه صرفاً بر اساس مجموعه‌ای ثابت از قوانین عمل کند، از داده‌های آموزشی الگوهای مربوط به رفتار سالم و مخرب را یاد می‌گیرد.

روش‌های یادگیری ماشین را می‌توان به سه گروه کلی تقسیم کرد: یادگیری نظارت‌شده، یادگیری بدون نظارت و یادگیری تقویتی. در یادگیری نظارت‌شده، داده‌های آموزشی دارای برچسب هستند؛ برای مثال مشخص است که یک فایل بدافزار است یا سالم. در یادگیری بدون نظارت، الگوریتم تلاش می‌کند الگوهای طبیعی و غیرطبیعی را بدون داشتن برچسب کامل کشف کند. یادگیری تقویتی نیز بر تعامل عامل هوشمند با محیط و یادگیری از نتایج اقدامات تمرکز دارد.

این قابلیت‌ها برای امنیت سایبری اهمیت ویژه‌ای دارند؛ زیرا بسیاری از حملات جدید فاقد امضای شناخته‌شده هستند. در چنین شرایطی، تشخیص ناهنجاری و تحلیل رفتاری می‌تواند مکمل روش‌های سنتی باشد.

۳. محدودیت روش‌های سنتی تشخیص حملات سایبری

روش‌های سنتی امنیت سایبری همچنان بخش مهمی از معماری دفاعی سازمان‌ها هستند. آنتی ویروس های مبتنی بر امضا، فایروال‌ها و سامانه‌های تشخیص نفوذ مبتنی بر قواعد، در برابر بسیاری از تهدیدات شناخته‌شده عملکرد مناسبی دارند.

با این حال، این روش‌ها با محدودیت‌هایی مواجه هستند. یکی از مهم‌ترین محدودیت‌ها وابستگی به شناخت قبلی تهدید است. اگر بدافزار یا الگوی حمله جدید باشد و امضای آن در پایگاه داده امنیتی وجود نداشته باشد، تشخیص آن دشوارتر خواهد بود.

از طرف دیگر، مهاجمان می‌توانند رفتار خود را تغییر دهند. برای مثال، یک حمله ممکن است به‌صورت آهسته و مرحله‌ای انجام شود تا از آستانه‌های تشخیص سامانه‌های سنتی عبور نکند.

همچنین تعداد زیاد هشدارها می‌تواند موجب «خستگی هشدار» در کارشناسان امنیتی شود. زمانی که متخصص امنیت با هزاران هشدار مواجه باشد، تشخیص هشدارهای واقعاً مهم دشوار خواهد شد. هوش مصنوعی می‌تواند با تحلیل و اولویت‌بندی هشدارها، این مشکل را کاهش دهد.

۴. کاربرد هوش مصنوعی در تشخیص نفوذ

یکی از مهم‌ترین کاربردهای هوش مصنوعی، تشخیص نفوذ در شبکه و سامانه‌های رایانه‌ای است.

سامانه تشخیص نفوذ مبتنی بر هوش مصنوعی می‌تواند داده‌هایی مانند آدرس‌های شبکه، پورت‌ها، پروتکل‌ها، حجم ترافیک، زمان ارتباط، تعداد درخواست‌ها و سایر ویژگی‌های ارتباطی را تحلیل کند. سپس با استفاده از مدل یادگیری ماشین، احتمال مخرب بودن یک ارتباط یا رفتار را تعیین می‌کند.

در روش‌های سنتی، معمولاً مجموعه‌ای از قواعد مشخص برای تشخیص حمله تعریف می‌شود؛ اما در روش‌های هوشمند، سیستم می‌تواند الگوی رفتار طبیعی شبکه را نیز یاد بگیرد. اگر رفتار جدید به شکل محسوسی از الگوی عادی فاصله داشته باشد، سیستم می‌تواند آن را به‌عنوان رفتار مشکوک علامت‌گذاری کند.

برای مثال، اگر یک حساب کاربری که معمولاً در ساعات اداری فعالیت می‌کند، نیمه‌شب از یک موقعیت جغرافیایی غیرمعمول وارد سامانه شده و حجم زیادی از فایل‌ها را دانلود کند، سیستم هوشمند می‌تواند این رفتار را به‌عنوان یک ناهنجاری شناسایی کند.

۵. کاربرد هوش مصنوعی در تشخیص بدافزار

بدافزار یکی از رایج ترین تهدیدات سایبری است و شامل ویروس‌ها، کرم‌ها، تروجان‌ها، جاسوس‌افزارها، باج‌افزارها و سایر نرم‌افزارهای مخرب می‌شود.

روش‌های سنتی تشخیص بدافزار اغلب بر بررسی امضای فایل متکی هستند. این روش در برابر نمونه‌های شناخته‌شده مؤثر است، اما مهاجمان می‌توانند با تغییر ساختار فایل یا استفاده از تکنیک‌های مبهم‌سازی، تشخیص بدافزار را دشوار کنند.

هوش مصنوعی امکان تحلیل ویژگی‌های ایستا و پویا را فراهم می‌کند. در تحلیل ایستا، ویژگی‌هایی مانند ساختار فایل، فراخوانی‌های موجود، کتابخانه‌ها و سایر مشخصات بررسی می‌شوند. در تحلیل پویا، رفتار برنامه هنگام اجرا مورد بررسی قرار می‌گیرد.

مدل‌های یادگیری ماشین می‌توانند با تحلیل این ویژگی‌ها، احتمال مخرب بودن یک فایل را تعیین کنند. این رویکرد به‌خصوص برای شناسایی گونه‌هایی از بدافزار که قبلاً مشاهده نشده‌اند اهمیت دارد.

۶. کاربرد هوش مصنوعی در مقابله با فیشینگ

فیشینگ یکی از مهم‌ترین روش‌های حمله سایبری است که در آن مهاجم تلاش می‌کند قربانی را فریب دهد تا اطلاعات حساس مانند نام کاربری، گذرواژه، اطلاعات بانکی یا سایر داده‌های محرمانه را افشا کند.

هوش مصنوعی می‌تواند پیام‌های الکترونیکی، محتوای وب‌سایت‌ها و الگوهای ارتباطی را برای شناسایی نشانه‌های فیشینگ تحلیل کند. الگوریتم‌ها می‌توانند ویژگی‌هایی مانند ساختار متن، نشانی اینترنتی، دامنه، محتوای پیام، رفتار فرستنده و تفاوت با الگوهای ارتباطی معمول را بررسی کنند.

پردازش زبان طبیعی نیز در این حوزه کاربرد دارد. با استفاده از این فناوری، سیستم می‌تواند متن پیام‌ها را تحلیل کرده و نشانه‌های مهندسی اجتماعی را شناسایی کند.

برای مثال، اگر پیام الکترونیکی با لحن اضطراری از کاربر بخواهد فوراً گذرواژه خود را تغییر دهد و لینک موجود در آن به دامنه‌ای مشکوک هدایت شود، سیستم هوشمند می‌تواند پیش از رسیدن پیام به کاربر، احتمال فیشینگ بودن آن را افزایش دهد.

۷. تشخیص حملات منع خدمت توزیع‌شده

حملات DDoS با ارسال حجم زیادی از درخواست‌ها به یک سرویس، تلاش می‌کنند منابع آن را مصرف و سرویس را از دسترس کاربران خارج کنند.

تشخیص این حملات صرفاً بر اساس یک آستانه ثابت همیشه مؤثر نیست؛ زیرا حجم طبیعی ترافیک در سامانه‌های مختلف متفاوت است. همچنین مهاجم می‌تواند حجم حمله را به‌گونه‌ای تنظیم کند که در نگاه اول شبیه افزایش طبیعی ترافیک باشد.

هوش مصنوعی می‌تواند الگوی ترافیک را در طول زمان یاد بگیرد و میان افزایش طبیعی و غیرطبیعی ترافیک تمایز ایجاد کند. مدل‌های تشخیص ناهنجاری می‌توانند تغییرات غیرمعمول در حجم درخواست‌ها، توزیع منابع و رفتار کاربران را شناسایی کنند.

بر اساس پژوهش‌های ENISA، روش‌های مبتنی بر هوش مصنوعی و یادگیری ماشین از جمله روش‌هایی هستند که برای مقابله با حملاتی مانند DDoS مورد بررسی قرار گرفته‌اند.

۸. شناسایی حملات باج افزاری

باج‌افزار یکی از تهدیدات مهم علیه سازمان‌ها است که معمولاً با رمزگذاری فایل‌ها یا اختلال در دسترسی به منابع، قربانی را با درخواست باج مواجه می‌کند.

یکی از مزیت‌های هوش مصنوعی در مقابله با باج‌افزار، امکان شناسایی رفتارهای غیرعادی قبل از تکمیل فرایند حمله است. برای مثال، اگر یک فرایند به‌طور ناگهانی شروع به دسترسی گسترده به تعداد زیادی فایل کند و الگوی رفتاری آن با فعالیت معمول تفاوت داشته باشد، سامانه امنیتی می‌تواند آن را بررسی کند.

این رویکرد به جای وابستگی کامل به نام یا امضای یک باج‌افزار خاص، بر رفتار آن تمرکز دارد و بنابراین می‌تواند در شناسایی گونه‌های جدید نیز مفید باشد.

۹. تحلیل رفتار کاربران و موجودیت‌ها

یکی از کاربردهای مهم هوش مصنوعی، تحلیل رفتار کاربران و موجودیت‌ها یا UEBA است.

در این روش، سامانه رفتار عادی کاربران، دستگاه‌ها و حساب‌های کاربری را یاد می‌گیرد. سپس هرگونه انحراف قابل توجه از این الگو را بررسی می‌کند.

برای مثال، اگر یک کارمند معمولاً به تعداد محدودی از فایل‌های سازمانی دسترسی داشته باشد، اما ناگهان حجم بسیار زیادی از اطلاعات را دانلود کند، سامانه می‌تواند این رفتار را مشکوک تشخیص دهد.

این قابلیت در شناسایی تهدیدات داخلی، سرقت حساب کاربری و سوءاستفاده از دسترسی‌های مجاز اهمیت دارد.

۱۰. پیش‌بینی و پیشگیری از حملات سایبری

نقش هوش مصنوعی تنها به تشخیص حملات پس از آغاز آن‌ها محدود نمی‌شود. یکی از اهداف مهم استفاده از هوش مصنوعی، حرکت از امنیت واکنشی به امنیت پیش‌نگر است.

در امنیت واکنشی، سازمان پس از وقوع حادثه واکنش نشان می‌دهد؛ اما در رویکرد پیش‌نگر، تلاش می‌شود نشانه‌های اولیه حمله قبل از وقوع خسارت جدی شناسایی شود.

هوش مصنوعی می‌تواند با تحلیل داده‌های تاریخی، اطلاعات تهدید، آسیب‌پذیری‌های موجود، الگوهای رفتاری مهاجمان و وضعیت دارایی‌های سازمان، نقاطی را که احتمال حمله به آن‌ها بیشتر است مشخص کند.

۱۱. کاربرد هوش مصنوعی در مدیریت آسیب‌پذیری‌ها

سازمان‌های بزرگ معمولاً با تعداد زیادی آسیب‌پذیری نرم‌افزاری مواجه هستند. همه آسیب‌پذیری‌ها از اهمیت یکسان برخوردار نیستند و رفع همزمان همه آن‌ها معمولاً امکان‌پذیر نیست.

هوش مصنوعی می‌تواند با تحلیل عواملی مانند شدت آسیب‌پذیری، میزان دسترسی، ارزش دارایی، وضعیت شبکه، سابقه حملات و احتمال سوءاستفاده، آسیب‌پذیری‌ها را اولویت‌بندی کند.

به این ترتیب، تیم امنیتی می‌تواند منابع محدود خود را ابتدا برای آسیب‌پذیری‌هایی اختصاص دهد که بیشترین ریسک را ایجاد می‌کنند.

این رویکرد باعث می‌شود مدیریت امنیت از یک فرایند صرفاً فنی به یک فرایند مبتنی بر ریسک تبدیل شود.

۱۲. خودکارسازی واکنش به رخدادهای امنیتی

یکی از مهم‌ترین مزایای هوش مصنوعی، امکان خودکارسازی بخشی از واکنش به رخدادها است.

برای نمونه، هنگامی که سامانه هوشمند احتمال یک حمله را بالا تشخیص می‌دهد، می‌توان اقداماتی مانند مسدودکردن یک ارتباط، قرنطینه‌کردن یک فایل، غیرفعال‌کردن موقت یک حساب کاربری یا ارسال هشدار با اولویت بالا را به‌صورت خودکار انجام داد.

این موضوع اهمیت زیادی دارد؛ زیرا سرعت حملات سایبری می‌تواند بسیار بیشتر از سرعت واکنش انسان باشد. در چنین شرایطی، خودکارسازی می‌تواند فاصله زمانی میان شناسایی و واکنش را کاهش دهد.

البته تصمیم‌های حساس نباید بدون کنترل انسانی به سیستم هوش مصنوعی واگذار شوند. به‌ویژه در زیرساخت‌های حیاتی، اقدامات خودکار باید دارای محدودیت، ثبت رویداد، امکان بازگشت و سازوکار تأیید انسانی باشند.

۱۳. نقش هوش مصنوعی در مراکز عملیات امنیت

مرکز عملیات امنیت یا SOC وظیفه پایش مستمر زیرساخت سازمان و پاسخ به رخدادهای امنیتی را بر عهده دارد.

یکی از مشکلات اصلی SOCها، حجم بسیار زیاد داده‌ها و هشدارها است. هوش مصنوعی می‌تواند داده‌های مختلف را از فایروال‌ها، سامانه‌های تشخیص نفوذ، سرورها، نقاط پایانی، سامانه‌های مدیریت رویداد و منابع اطلاعات تهدید دریافت و با یکدیگر مرتبط کند.

برای مثال، مشاهده یک ورود غیرمعمول به حساب کاربری به‌تنهایی ممکن است یک رخداد عادی باشد؛ اما اگر همزمان با آن، ارتباط با یک نشانی اینترنتی مشکوک و اجرای یک فایل ناشناخته نیز مشاهده شود، احتمال وقوع حمله افزایش می‌یابد.

بنابراین هوش مصنوعی می‌تواند به جای بررسی جداگانه رویدادها، ارتباط میان آن‌ها را کشف کند.

۱۴. هوش مصنوعی و تشخیص حملات ناشناخته

یکی از مهم‌ترین مزایای رویکردهای هوشمند، ظرفیت آن‌ها برای کشف رفتارهای ناشناخته است.

در حملات روز-صفر، مهاجم از آسیب‌پذیری‌ای استفاده می‌کند که ممکن است هنوز برای عموم شناخته نشده باشد. در این شرایط، روش‌های مبتنی بر امضا ممکن است نتوانند حمله را تشخیص دهند.

روش‌های تشخیص ناهنجاری می‌توانند در این زمینه مفید باشند؛ زیرا الزاماً به شناخت قبلی نمونه حمله نیاز ندارند و به جای آن، رفتار غیرعادی را شناسایی می‌کنند.

بااین‌حال، تشخیص ناهنجاری نیز محدودیت دارد. هر رفتار غیرعادی الزاماً مخرب نیست و همین موضوع می‌تواند تعداد هشدارهای کاذب را افزایش دهد.

۱۵. چالش‌های استفاده از هوش مصنوعی در امنیت سایبری

با وجود ظرفیت بالای هوش مصنوعی، استفاده از این فناوری بدون چالش نیست.

15-1. کیفیت داده

مدل های یادگیری ماشین به داده وابسته هستند. اگر داده‌های آموزشی ناقص، قدیمی، نامتوازن یا دارای خطا باشند، مدل نیز ممکن است عملکرد مناسبی نداشته باشد.

15-2. هشدارهای کاذب

اگر مدل بیش از اندازه حساس باشد، تعداد زیادی رفتار سالم را نیز به‌عنوان حمله شناسایی خواهد کرد. افزایش هشدارهای کاذب می‌تواند باعث کاهش اعتماد کارشناسان به سامانه شود.

15-3. توضیح پذیری

در بسیاری از مدل‌های پیچیده یادگیری عمیق مشخص نیست چرا مدل یک رویداد را مخرب تشخیص داده است. در محیطهای حساس، کارشناسان امنیتی باید بتوانند منطق تصمیم سامانه را تا حد قابل قبول درک کنند.

15-4. حملات خصمانه علیه هوش مصنوعی

یکی از مهم‌ترین چالش‌ها، حمله به خود مدل هوش مصنوعی است. مهاجم می‌تواند با ایجاد تغییرات هدفمند در ورودی‌ها تلاش کند مدل را فریب دهد.

15-5. مسموم سازی داده

در حملات مسموم سازی، مهاجم تلاش می‌کند داده‌های آموزشی مدل را دستکاری کند. اگر این داده‌ها وارد فرایند آموزش شوند، ممکن است مدل در آینده تصمیم‌های نادرست بگیرد.

15-6. حریم خصوصی

استفاده از داده‌های کاربران برای تحلیل رفتاری می‌تواند نگرانی‌های مربوط به حریم خصوصی ایجاد کند. بنابراین سازمان‌ها باید میان امنیت، نیاز به داده و حقوق کاربران توازن برقرار کنند.

۱۶. سوءاستفاده مهاجمان از هوش مصنوعی

هوش مصنوعی تنها در اختیار مدافعان نیست. مهاجمان نیز می‌توانند از آن برای افزایش کارایی حملات خود استفاده کنند.

از جمله کاربردهای مخرب احتمالی می‌توان به تولید پیام‌های فیشینگ متقاعدکننده‌تر، خودکارسازی برخی فرایندهای شناسایی اهداف، تولید محتوای جعلی، بهبود مهندسی اجتماعی و افزایش مقیاس حملات اشاره کرد.

این موضوع نشان می‌دهد که توسعه دفاع مبتنی بر هوش مصنوعی باید همزمان با توسعه روش‌های محافظت از خود سامانه‌های هوشمند انجام شود.

۱۷. راهکار پیشنهادی برای استفاده مؤثر از هوش مصنوعی

برای استفاده موفق از هوش مصنوعی در امنیت سایبری، سازمان‌ها باید از رویکردی چندلایه استفاده کنند.

نخست، باید دارایی‌های سازمان شناسایی و طبقه‌بندی شوند. سپس داده‌های امنیتی از منابع مختلف جمع‌آوری و استانداردسازی شوند. پس از آن می‌توان مدل‌های مناسب را بر اساس نوع تهدید و نیاز سازمان انتخاب کرد.

در مرحله بعد، مدل باید با داده‌های واقعی و معتبر ارزیابی شود. معیارهایی مانند دقت، نرخ تشخیص، نرخ هشدار کاذب، زمان تشخیص و زمان واکنش باید مورد توجه قرار گیرند.

همچنین باید یک سازوکار نظارت انسانی وجود داشته باشد. تصمیم‌های مهم امنیتی نباید صرفاً به خروجی یک مدل واگذار شوند.

به‌روزرسانی مدل نیز اهمیت دارد؛ زیرا رفتار مهاجمان دائماً تغییر می‌کند و مدل آموزش‌دیده بر اساس داده‌های قدیمی ممکن است به مرور کارایی خود را از دست بدهد.

۱۸. وضعیت و اهمیت موضوع در ایران

افزایش استفاده از خدمات دیجیتال در ایران نیز ضرورت توجه به امنیت سایبری و فناوری‌های هوشمند را افزایش داده است. سازمان‌های دولتی، بانک‌ها، شرکت‌های فناوری، مراکز دانشگاهی، صنایع و زیرساخت‌های حیاتی به سامانه‌های دیجیتال وابستگی زیادی دارند.

در سند ملی هوش مصنوعی جمهوری اسلامی ایران نیز توسعه و کاربرد هوش مصنوعی به‌عنوان یک موضوع راهبردی مورد توجه قرار گرفته است. این سند در سال ۱۴۰۳ به تصویب شورای عالی انقلاب فرهنگی رسیده است.

از منظر امنیت سایبری، استفاده از هوش مصنوعی می‌تواند در حوزه‌هایی مانند پایش شبکه‌های سازمانی، تشخیص بدافزار، تحلیل رخدادهای امنیتی، مقابله با فیشینگ، شناسایی رفتارهای غیرعادی و مدیریت آسیب‌پذیری‌ها مورد توجه قرار گیرد.

با این حال، توسعه این حوزه نیازمند ایجاد مجموعه‌داده‌های بومی، تربیت نیروی انسانی متخصص، توسعه آزمایشگاه‌های امنیت سایبری و هوش مصنوعی، استانداردسازی روش‌های ارزیابی و توجه به امنیت خود مدل‌های هوش مصنوعی است.

۱۹. آینده هوش مصنوعی در امنیت سایبری

آینده امنیت سایبری احتمالاً به سمت سامانه‌هایی حرکت خواهد کرد که توانایی بیشتری در تحلیل خودکار، پیش‌بینی تهدید، تصمیم‌گیری و واکنش سریع دارند.

مدل‌های زبانی بزرگ نیز می‌توانند در تحلیل گزارش‌های امنیتی، خلاصه‌سازی رخدادها، جست‌وجوی اطلاعات تهدید، کمک به کارشناسان SOC و تولید توضیحات قابل فهم درباره هشدارها کاربرد داشته باشند.

بااین‌حال، افزایش خودکارسازی به معنای حذف کامل نقش انسان نیست. برعکس، هرچه سامانه‌ها پیچیده‌تر شوند، نیاز به متخصصانی که بتوانند خروجی مدل‌ها را ارزیابی، خطاهای آن‌ها را شناسایی و تصمیم‌های حساس را کنترل کنند، افزایش خواهد یافت.

بنابراین آینده مطلوب امنیت سایبری، ترکیبی از هوش مصنوعی و تخصص انسانی است؛ به‌گونه‌ای که ماشین وظایف حجیم، تکراری و زمان‌بر را انجام دهد و انسان بر تصمیم‌گیری‌های حساس، راهبردی و اخلاقی نظارت داشته باشد.

۲۰. نتیجه‌گیری

تحول ماهیت حملات سایبری موجب شده است روش‌های سنتی امنیت اطلاعات به‌تنهایی پاسخگوی نیازهای محیط دیجیتال امروز نباشند. حملات جدید می‌توانند سریع، چندمرحله‌ای، ناشناخته و سازگار با شرایط محیطی باشند و حجم بالای داده‌های امنیتی نیز تحلیل دستی آن‌ها را دشوار می‌کند.

هوش مصنوعی ظرفیت قابل توجهی برای حل بخشی از این مشکلات دارد. یادگیری ماشین و یادگیری عمیق می‌توانند برای تشخیص نفوذ، کشف ناهنجاری، شناسایی بدافزار، تشخیص فیشینگ، مقابله با DDoS، شناسایی باج‌افزار، تحلیل رفتار کاربران، مدیریت آسیب‌پذیری‌ها و خودکارسازی واکنش به رخدادها مورد استفاده قرار گیرند.

مهم‌ترین مزیت هوش مصنوعی، توانایی آن در تحلیل سریع حجم زیادی از داده و شناسایی الگوهایی است که ممکن است برای انسان دشوار باشند. بااین‌حال، این فناوری راه‌حل کامل و بدون نقص نیست. کیفیت داده، هشدارهای کاذب، توضیح‌پذیری، حریم خصوصی، هزینه پیاده‌سازی و حملات خصمانه علیه مدل‌ها از جمله چالش‌های اساسی آن هستند.

از سوی دیگر، مهاجمان نیز می‌توانند از هوش مصنوعی برای افزایش کارایی حملات استفاده کنند. بنابراین امنیت سایبری در عصر هوش مصنوعی تنها به معنای «استفاده از هوش مصنوعی برای دفاع» نیست؛ بلکه باید شامل «محافظت از هوش مصنوعی» و «مدیریت سوءاستفاده احتمالی از هوش مصنوعی» نیز باشد.

در مجموع، می‌توان نتیجه گرفت که بهترین رویکرد، استفاده از هوش مصنوعی به‌عنوان یکی از اجزای یک معماری دفاعی چندلایه است. این معماری باید از ترکیب فناوری‌هایی مانند فایروال، سامانه‌های تشخیص و پیشگیری از نفوذ، مدیریت رخدادهای امنیتی، اطلاعات تهدید، کنترل دسترسی، آموزش کاربران و سامانه‌های هوشمند تشکیل شود.

بنابراین، آینده امنیت سایبری نه در جایگزینی انسان با ماشین، بلکه در ایجاد همکاری مؤثر میان متخصصان انسانی و سامانه‌های هوشمند نهفته است. سازمان‌هایی که بتوانند این همکاری را بر پایه داده‌های باکیفیت، مدل‌های قابل اعتماد، نظارت انسانی و اصول امنیتی مناسب ایجاد کنند، آمادگی بیشتری برای شناسایی و پیشگیری از حملات سایبری آینده خواهند داشت.

منابع

۱. شورای عالی انقلاب فرهنگی. (۱۴۰۳). سند ملی هوش مصنوعی جمهوری اسلامی ایران. مصوب شورای عالی انقلاب فرهنگی.

۲. شورای عالی انقلاب فرهنگی. (۱۴۰۲). مصوبه نهایی سازی و تصویب سند ملی هوش مصنوعی جمهوری اسلامی ایران.

۳. مؤسسه ملی استاندارد و فناوری ایالات متحده آمریکا (NIST) راهنمای سامانه‌های تشخیص و پیشگیری از نفوذ. (800-94SP) ترجمه و استفاده با هدف پژوهشی.

۴. آژانس اتحادیه اروپا برای امنیت سایبری (ENISA) هوش مصنوعی و چالش‌های امنیت سایبری. گزارش پژوهشی درباره رابطه میان هوش مصنوعی، امنیت سایبری و تهدیدات نوظهور.

۵. آژانس اتحادیه اروپا برای امنیت سایبری (ENISA) پژوهش درباره هوش مصنوعی و امنیت سایبری. گزارش پژوهشی درباره کاربردهای هوش مصنوعی در تشخیص و پاسخ به تهدیدات.

6. Scarfone, K. A., & Mell, P. (2007). Guide to Intrusion Detection and Prevention Systems (IDPS). NIST Special Publication 800-94. National Institute of Standards and Technology .

7. European Union Agency for Cybersecurity (ENISA). (2023). AI Good Cybersecurity Practices: A Multilayer Framework for Good Cybersecurity Practices for AI .

8. European Union Agency for Cybersecurity (ENISA). (2020). Artificial Intelligence Cybersecurity Challenges .

9. European Union Agency for Cybersecurity (ENISA). (2023). Artificial Intelligence and Cybersecurity Research .

10. Vassilev, A., Oprea, A., Fordyce, A., Anderson, H., Davies, X., & Hamin, M. (2025). Adversarial Machine Learning: A Taxonomy and Terminology of Attacks and Mitigations. NIST AI 100-2E .2025

11. National Institute of Standards and Technology (NIST). (2024). Artificial Intelligence Technology and Cybersecurity Resources .

12. National Institute of Standards and Technology (NIST). Intrusion Detection. Computer Security Resource Center (CSRC .(

13. National Institute of Standards and Technology (NIST). Protecting Controlled Unclassified Information in Nonfederal Systems and Organizations. NIST SP 800-171.

14. European Union Agency for Cybersecurity (ENISA). (2022). Research and Innovation Brief: Artificial Intelligence and Cybersecurity.