# هوش مصنوعی در حملات سایبری

**شهلا منیعی فرد**
دانشجوی دکتری کامپیوتر گرایش هوش مصنوعی دانشگاه ازاد تهران

## چکیده
پیشرفت سریع هوش مصنوعی (AI) تأثیر قابل‌توجهی بر حوزه‌های مختلف، از جمله امنیت سایبری، داشته است. این مطالعه ابعاد نظری استفاده از هوش مصنوعی در حملات سایبری را بررسی می‌کند و نقش آن را در خودکارسازی استراتژی‌های پیچیده حمله، افزایش اثربخشی روش‌های مهاجم و اجرای کمپین‌های پیچیده مهندسی اجتماعی برجسته می‌کند. با استفاده از الگوریتم‌های یادگیری ماشین، مهاجمان می‌توانند آسیب‌پذیری‌ها را بهره‌برداری کنند، از سیستم‌های تشخیص عبور کنند و حملات بسیار هدفمند و سازگار انجام دهند که تهدیدی جدی برای زیرساخت‌های دیجیتال به شمار می‌رود.این مطالعه بر ضرورت استفاده از راهکارهای مبتنی بر هوش مصنوعی، از جمله تحلیل پیش‌بینی‌کننده، شناسایی ناهنجاری‌ها و هوش تهدید مشارکتی، برای مقابله پیشگیرانه با این چالش‌ها تأکید می‌کند.علاوه بر این، این مطالعه مسیرهای پژوهشی مهمی از جمله توسعه مکانیزم‌های دفاعی پیشرفته، همکاری‌های بین‌رشته‌ای و ایجاد دستورالعمل‌های اخلاقی قوی را ترسیم می‌کند. این تحلیل جامع با هدف ارائه درکی عمیق‌تر از نقش هوش مصنوعی در حملات سایبری و ترویج رویکردهای نوآورانه برای حفاظت از یکپارچگی اکوسیستم‌های دیجیتال انجام شده است.

**واژگان کلیدی:** تهدیدات امنیتی ، یادگیری ماشین (ML) ،حملات سایبری ، هوش مصنوعی (AI) ، هوش مصنوعی متخاصم.

**مقدمه**

هوش مصنوعی (AI)، به عنوان یک فناوری تحول‌آفرین، پیشرفت‌های چشمگیری در بازتعریف چشم‌انداز فناوری داشته است. هوش مصنوعی شامل زیرشاخه‌های متعددی مانند یادگیری ماشین (ML)، پردازش زبان طبیعی (NLP)، بینایی کامپیوتری و رباتیک است که هرکدام به شیوه‌ای منحصربه‌فرد به حل مسائل پیچیده کمک می‌کنند. در حوزه امنیت سایبری، هوش مصنوعی به عنوان یک راه‌حل محوری ظاهر شده است و امکان توسعه ابزارها و سیستم‌هایی را فراهم کرده که قادر به شناسایی، پیشگیری و پاسخ به تهدیدات سایبری در زمان واقعی هستند.هسته اصلی قدرت هوش مصنوعی در توانایی آن برای پردازش و تحلیل حجم عظیمی از داده‌ها با سرعتی فراتر از توانایی انسان‌ها نهفته است. روش‌های سنتی امنیت سایبری عمدتاً به تحلیل دستی و سیستم‌های مبتنی بر قوانین از پیش تعریف‌شده متکی هستند که در مقابله با حجم و پیچیدگی تهدیدات مدرن با مشکل مواجه می‌شوند. از طریق یادگیری ماشین، هوش مصنوعی سیستم‌های پویایی معرفی کرده است که به الگوها و تهدیدات نوظهور واکنش نشان می‌دهند. برای مثال، سیستم‌های شناسایی ناهنجاری مبتنی بر هوش مصنوعی می‌توانند رفتارهای غیرعادی را در یک شبکه شناسایی کنند، در حالی که تحلیل پیش‌بینی‌کننده به سازمان‌ها کمک می‌کند تا آسیب‌پذیری‌ها را پیش از بهره‌برداری رفع کنند.با این حال، ویژگی‌هایی که هوش مصنوعی را به یک ابزار دفاعی قدرتمند تبدیل کرده‌اند، آن را به دارایی جذابی برای مهاجمان نیز بدل کرده‌اند. با استفاده از هوش مصنوعی، عوامل مخرب می‌توانند استراتژی‌های پیچیده حمله را خودکار کنند و در نتیجه آن‌ها را کارآمدتر و مقیاس‌پذیرتر سازند. برای مثال، کمپین‌های فیشینگ خودکار که توسط پردازش زبان طبیعی تقویت شده‌اند، می‌توانند پیام‌های بسیار متقاعدکننده و شخصی‌سازی‌شده‌ای را در مقیاس وسیع ایجاد کنند که حتی محتاط‌ترین کاربران را فریب دهند. به‌طور مشابه، بدافزارهای تقویت‌شده با هوش مصنوعی می‌توانند رفتار خود را در زمان واقعی تغییر دهند تا از شناسایی توسط سیستم‌های آنتی‌ویروس سنتی اجتناب کنند. ماهیت دوگانه هوش مصنوعی ،توانایی آن در تقویت دفاع و همزمان افزایش حملات،چالش‌های قابل‌توجهی را برای متخصصان امنیت سایبری ایجاد می‌کند.

**روش تحقیق**

ادغام هوش مصنوعی در حملات سایبری، تغییری اساسی در ماهیت تهدیدات دیجیتال ایجاد کرده و چالش‌های حیاتی متعددی را به همراه دارد.

- **دور زدن مکانیزم های امنیتی سنتی:** هوش مصنوعی به مهاجمان امکان می‌دهد حملات تطبیق‌پذیر طراحی کنند و از سیستم‌های دفاعی عبور کنند. تکنیک‌هایی مانند یادگیری ماشین خصمانه می‌توانند ورودی‌ها را دستکاری کرده و سیستم‌های هوش مصنوعی را فریب دهند. [14]
- **خودکارسازی و مقیاس‌پذیری:** حملات مبتنی بر هوش مصنوعی با سرعت و مقیاسی بسیار بالا انجام می‌شوند، مانند اسکن شبکه‌ها برای آسیب‌پذیری‌ها یا اجرای حملات DDoS، که مقابله با آن‌ها برای تیم‌های سنتی امنیت دشوار است. [15]
- **نبود قوانین و نظارت اخلاقی:** نبود قوانین و نظارت کافی، محیطی آزاد برای توسعه ابزارهای حمله مبتنی بر هوش مصنوعی فراهم کرده و ردیابی و کاهش تهدیدات را پیچیده‌تر کرده است. [16]
- **تضعیف اعتماد:** استفاده از فناوری‌هایی مانند دیپ‌فیک منجر به انتشار اطلاعات نادرست، کاهش اعتماد به سیستم‌های دیجیتال و آسیب‌های اجتماعی یا مالی می‌شود.[13]

**اهداف**

هدف اصلی این مطالعه، تحلیل تلاقی هوش مصنوعی (AI) و حملات سایبری است و بر فرصت‌ها و چالش‌های مرتبط با این فناوری دوگانه تمرکز دارد. برای دستیابی به این هدف، مطالعه بر اهداف زیر متمرکز شده است. [7]

- **تحلیل وضعیت فعلی:** بررسی کاربردهای پیشرفته هوش مصنوعی در امنیت سایبری، از جمله نقش آن به‌عنوان ابزار دفاعی یا سلاح تهاجمی، همراه با مثال‌هایی مانند استفاده از هوش مصنوعی خصمانه برای عبور از سیستم‌های دفاعی و تقویت بدافزارها. [5]
- **شناسایی چالش‌ها:** این مطالعه چالش‌های فنی، اخلاقی، و نظارتی هوش مصنوعی، از جمله محدودیت‌ها در شناسایی تهدیدات، نگرانی‌های حریم خصوصی، و نبود پروتکل‌های استاندارد را بررسی می‌کند. [18]
- **پیشنهاد مسیرهای تحقیقاتی آینده:** این مطالعه استراتژی‌هایی مانند توسعه سیستم‌های دفاعی هوش مصنوعی، همکاری‌های بین‌رشته‌ای برای تدوین دستورالعمل‌های اخلاقی، و توصیه‌های سیاست‌گذاری برای استفاده مسئولانه از هوش مصنوعی ارائه می‌دهد. [۱۹]

با پرداختن به این اهداف، این مطالعه قصد دارد به درک عمیق‌تری از نقش هوش مصنوعی در امنیت سایبری کمک کند و پایه‌ای برای تحقیقات آینده و راه‌حل‌های عملی برای مقابله با تهدیدات ناشی از هوش مصنوعی ارائه دهد. [۴]

## کاربردهای هوش مصنوعی در حملات سایبری

- **خودکارسازی حملات:** هوش مصنوعی روش اجرای حملات سایبری را با معرفی قابلیت‌های پیشرفته خودکارسازی متحول کرده است. این خودکارسازی حملات را کارآمدتر، مقیاس‌پذیرتر و تطبیق‌پذیرتر کرده و به مهاجمان امکان می‌دهد تا با کمترین دخالت انسانی سیستم‌ها را هدف قرار دهند. [۱۱]

- **اسکن آسیب‌پذیری‌ها:** ابزارهای هوش مصنوعی مانند Shodan و Nessus با خودکارسازی اسکن آسیب‌پذیری‌ها از طریق یادگیری ماشین، شناسایی نقاط ضعف در فایروال‌ها، اپلیکیشن‌های وب و سرورها را بهبود می‌بخشند و رفتار اسکن را براساس مکانیزم‌های دفاعی تطبیق می‌دهند. [27,20]

- **بهره‌برداری خودکار:** هوش مصنوعی به مهاجمان امکان می‌دهد بهره‌برداری از آسیب‌پذیری‌ها را خودکار کرده و بارهای حمله را به‌صورت پویا تنظیم کنند. ابزارهای مبتنی بر هوش مصنوعی می‌توانند آسیب‌پذیری‌ها مانند سرریز بافر را شناسایی کرده و شل‌کدهایی طراحی کنند تا دستورات دلخواه را اجرا کنند. [۶]

- **بدافزار پلی‌مورفیک:** هوش مصنوعی توسعه بدافزارهای پلی‌مورفیک را ممکن ساخته که ساختار خود را تغییر داده و از سیستم‌های آنتی‌ویروس فرار می‌کنند. این بدافزارها با استفاده از یادگیری تقویتی، عملکرد خود را بر اساس موفقیت در دور زدن سیستم‌های امنیتی تطبیق می‌دهند. [28,11]

- **فیشینگ در مقیاس وسیع:** با استفاده از پردازش زبان طبیعی (NLP) مبتنی بر هوش مصنوعی، مهاجمان می‌توانند ایمیل‌های فیشینگ شخصی‌سازی‌شده و واقع‌گرایانه‌ای تولید کنند که لحن و سبک ارتباطات قانونی را تقلید می‌کنند. این ایمیل‌ها از توانایی هوش مصنوعی در تحلیل داده‌های عمومی در شبکه‌های اجتماعی و وب‌سایت‌ها بهره می‌برند. [۲۱]

- **حملات خصمانه:** حملات خصمانه بر بهره‌برداری از نقاط ضعف موجود در مدل‌های یادگیری ماشین (ML) و هوش مصنوعی تمرکز دارند تا یکپارچگی و قابلیت اطمینان آن‌ها را به خطر بیندازند. این حملات به دلیل استفاده روزافزون از سیستم‌های هوش مصنوعی در کاربردهای حساس مانند تشخیص چهره، شناسایی نفوذ و پیشگیری از تقلب، نگرانی فزاینده‌ای ایجاد کرده‌اند. [۲۰]

- **حملات اجتنابی (Evasion Attacks):** در حملات اجتنابی، مهاجمان داده‌ها را دستکاری می‌کنند تا سیستم‌های هوش مصنوعی را فریب دهند. برای مثال، افزودن نویز به تصویر می‌تواند باعث اشتباه در شناسایی چهره‌ها شود. این روش در خودروهای خودران نیز باعث تفسیر اشتباه علائم راهنمایی و رانندگی و تصادفات می‌شود. [۲۲]

- **حملات مسموم‌سازی (Poisoning Attacks):** حملات مسموم‌سازی با وارد کردن داده‌های مخرب به داده‌های آموزشی، فرآیند یادگیری مدل‌های هوش مصنوعی را مختل کرده و رفتار آن‌ها را در شرایط خاص تغییر می‌دهند، مانند طبقه‌بندی اشتباه هرزنامه به‌عنوان پیام معتبر. [۲۳]

- **حملات تروجانی (Trojan Attacks):** در حملات تروجانی یا پشتی، مهاجمان محرک‌های مخفی را در یک مدل هوش مصنوعی در حین توسعه جاسازی می‌کنند. این محرک‌ها در شرایط خاص، رفتارهای مخرب را فعال می‌کنند. به‌عنوان مثال، یک مدل هوش مصنوعی آلوده به تروجان که در پهپادهای خودمختار استفاده می‌شود، ممکن است به‌طور عادی عمل کند اما با مشاهده یک الگوی بصری خاص به سلاح تبدیل شود. [۲۴]

- **استنتاج و استخراج مدل (Model Inference and Extraction):** مهاجمان با استفاده از پرس‌وجوهای مکرر به API مدل‌های هوش مصنوعی، از تکنیک استخراج مدل برای افشای معماری یا داده‌های حساس استفاده می‌کنند. این روش می‌تواند الگوریتم‌ها یا الگوهایی را شناسایی کند که به دور زدن شناسایی، مانند سیستم‌های مقابله با تقلب، کمک می‌کند. [۲۵]

- **مهندسی اجتماعی با استفاده از هوش مصنوعی:** مهندسی اجتماعی همچنان یکی از مؤثرترین روش‌ها برای نفوذ به سیستم‌ها از طریق بهره‌برداری از آسیب‌پذیری‌های انسانی به‌جای نقص‌های فنی است. با ادغام هوش مصنوعی، این حملات بسیار پیچیده‌تر، مقیاس‌پذیرتر و شخصی‌تر شده‌اند. [۲۶]

- **فناوری دیپ‌فیک (Deepfake Technology):** دیپ‌فیک‌ها، که توسط الگوریتم‌های هوش مصنوعی مانند GANs ساخته می‌شوند، به مهاجمان اجازه می‌دهند محتوای صوتی، تصویری و ویدیویی جعلی تولید کنند. این تکنیک‌ها در کلاهبرداری‌هایی مانند تقلید صدای مدیرعامل برای تأیید تراکنش‌های مالی جعلی و فیشینگ صوتی (وایشینگ) مورد استفاده قرار می‌گیرند. [۲۷]

- **تحلیل رفتاری مبتنی بر هوش مصنوعی:** سیستم‌های هوش مصنوعی می‌توانند با تحلیل فعالیت‌های آنلاین و پست‌های شبکه‌های اجتماعی یک هدف، پروفایل رفتاری دقیقی بسازند. این تحلیل به مهاجمان امکان می‌دهد تا کمپین‌های فیشینگ هدفمند (Spear-phishing) طراحی کنند، مانند ارسال ایمیل‌های مخرب با استفاده از اطلاعات اخیر هدف. [۲۸]

- **چت‌بات‌ها برای فیشینگ:** چت‌بات‌های مبتنی بر هوش مصنوعی می‌توانند مکالمات انسانی را شبیه‌سازی کنند و کاربران را فریب دهند تا اطلاعات حساس را به اشتراک بگذارند یا فایل‌های مخرب دانلود کنند. این چت‌بات‌ها با پردازش زبان طبیعی (NLP) توانایی پاسخ‌دهی دقیق به ورودی‌های متنی کاربر را دارند. [۲۹]

- **تحلیل احساسات برای مهندسی اجتماعی:** ابزارهای هوش مصنوعی می‌توانند لحن احساسی مکالمات را تحلیل کرده و موثرترین روش برای دستکاری هدف را شناسایی کنند. به‌عنوان مثال، یک مهاجم می‌تواند از هوش مصنوعی برای سوءاستفاده از ناامیدی یک کاربر و ظاهر شدن به‌عنوان نماینده مفید استفاده کند، که حملات مهندسی اجتماعی را موثرتر و مرز بین فریب انسانی و ماشینی را مبهم می‌سازد. [۳۰]

### مثال‌ها و مطالعات موردی

۱. **کمپین‌های بدافزاری Emotet**: Emotet از هوش مصنوعی برای ساخت ایمیل‌های فیشینگ شخصی‌سازی‌شده استفاده می‌کند که شبیه پاسخ‌های واقعی هستند و تعامل کاربران را افزایش می‌دهند. [۳۱]

۲. **DeepLocker توسط IBM**: DeepLocker از تشخیص چهره برای شناسایی قربانیان و ارسال بدافزار فقط زمانی که هدف شناسایی می‌شود، استفاده می‌کند، که خطر شناسایی و خسارات جانبی را کاهش می‌دهد. [۳۲]

۳. **حملات DDoS مبتنی بر هوش مصنوعی**: در یک مطالعه موردی ۲۰۲۱، مهاجمان با استفاده از هوش مصنوعی ترافیک حمله DDoS را به‌گونه‌ای تنظیم کردند که رفتار کاربران واقعی را شبیه‌سازی کرده و از استراتژی‌های کاهش آسیب فرار کردند. [۱۲]

٤. **کلاه‌برداری با کمک Deepfake**: مهاجمان با استفاده از فناوری Deepfake صدای مدیرعامل را جعل کردند و کارکنان را قانع کردند تا ۲۴۳,۰۰۰ دلار به حساب تقلبی منتقل کنند. [۳۳]

۵. **بدافزار با هوش مصنوعی (Ransomware)**: بدافزارهای رمزنگاری‌شده با هوش مصنوعی مانند "Egregor" از الگوریتم‌های انطباقی برای فرار از شناسایی استفاده می‌کنند. این بدافزار به‌طور پویا فایل‌ها را بر اساس ارزش آن‌ها برای سازمان هدف انتخاب می‌کند و فشار به پرداخت باج را افزایش می‌دهد. [۳۴]

این مثال‌ها تغییرات و تحول تهدیدات سایبری را نشان می‌دهند که با در دسترس بودن روزافزون فناوری‌های هوش مصنوعی، شدت یافته است. مقابله با این تهدیدات نیازمند رویکردی چندوجهی است که ترکیبی از حفاظت‌های فنی، ملاحظات اخلاقی و چارچوب‌های قانونی را در بر گیرد.

## فرضیات

### چالش‌ها و نگرانی‌های اخلاقی

- **چالش‌های شناسایی:** شناسایی حملات سایبری مبتنی بر هوش مصنوعی به دلیل طبیعت پویا و انطباق‌پذیر تهدیدات دشوار است. پیچیدگی و مقیاس‌پذیری حملات هوش مصنوعی باعث می‌شود که سیستم‌های دفاعی سنتی در بسیاری از موارد ناکارآمد شوند. [۳۵]
- **رفتار پویا و گریز از شناسایی:** حملات مبتنی بر هوش مصنوعی به‌طور مداوم تاکتیک‌های خود را تغییر می‌دهند تا از سیستم‌های شناسایی مانند آنتی‌ویروس مبتنی بر امضا فرار کنند. بدافزارهای پلی‌مورفیک و یادگیری ماشین خصمانه مثال‌هایی از این نوع حملات هستند. [۳۶]
- **حجم و سرعت بالا:** سیستم‌های هوش مصنوعی قادر به اجرای حملات با مقیاس و سرعتی فراتر از توانایی انسان‌ها هستند. این حملات می‌توانند میلیون‌ها کاربر را هدف قرار دهند و با محتوای شخصی‌سازی‌شده مواجه کنند. [47,36]
- **تکنیک‌های پیچیده پنهان‌سازی:** مهاجمان می‌توانند از مدل‌های تولیدی مانند GANs برای شبیه‌سازی ترافیک عادی استفاده کنند تا سیستم‌های تشخیص نفوذ را فریب دهند و بین فعالیت‌های مشروع و مخرب تمایز قائل نشوند. [49,37]
- **عدم قابلیت توضیح در دفاع‌های مبتنی بر هوش مصنوعی:** سیستم‌های دفاعی مبتنی بر هوش مصنوعی ممکن است به دلیل "جعبه سیاه" بودنشان نتوانند دلایل تصمیم‌گیری‌ها را ارائه دهند. این مشکل می‌تواند منجر به نتایج مثبت کاذب یا از دست دادن تهدیدات واقعی شود. [۱۰،۵۰]

## مدلسازی

### مقابله با چالش‌ها با استفاده از هوش مصنوعی قابل توضیح (XAI):

برای مقابله با چالش‌های ناشی از حملات سایبری مبتنی بر هوش مصنوعی، پژوهشگران و متخصصان امنیت سایبری به ادغام هوش مصنوعی قابل توضیح (XAI) در سیستم‌های دفاعی تأکید دارند. هدف XAI این است که فرآیندهای تصمیم‌گیری سیستم‌های هوش مصنوعی شفاف باشد و انسان‌ها بتوانند دلایل پشت تصمیمات آن‌ها را درک کنند. با ایجاد مدل‌های هوش مصنوعی که دلایل واضحی برای پیش‌بینی‌های خود ارائه می‌دهند، مدافعان می‌توانند بهتر خطرات امنیتی را ارزیابی کنند و اقدامات مناسب برای کاهش تهدیدات احتمالی انجام دهند. این شفافیت به ایجاد اعتماد در سیستم‌های دفاعی مبتنی بر هوش مصنوعی کمک می‌کند و نگرانی‌ها درباره مدل‌های "جعبه سیاه" را برطرف می‌سازد. [۳۸]

## مسائل اخلاقی

استفاده روزافزون از هوش مصنوعی در حملات سایبری نگرانی‌های اخلاقی جدی ایجاد می‌کند که باید برای جلوگیری از سوءاستفاده و حفاظت از افراد و جامعه مورد توجه قرار گیرد. این نگرانی‌ها جنبه‌های مختلفی دارند که هم به مهاجمانی که از هوش مصنوعی برای مقاصد مخرب استفاده می‌کنند و هم به سازمان‌هایی که باید از خود در برابر این تهدیدات دفاع کنند، مربوط می‌شوند. [۳۷]

- **معضل استفاده دوگانه:** مسئله اخلاقی استفاده دوگانه از هوش مصنوعی مطرح است، چرا که این فناوری می‌تواند هم برای اهداف مفید مانند بهبود مراقبت‌های بهداشتی و شناسایی تقلب استفاده شود و هم برای مقاصد مضر مانند نفوذ به سیستم‌های حساس یا انجام حملات سایبری. این چالش نیازمند مقررات دقیق در طراحی و پیاده‌سازی است. [55,39]
- **هدف قرار دادن جمعیت‌های آسیب‌پذیر:** حملات سایبری مبتنی بر هوش مصنوعی اغلب جمعیت‌های آسیب‌پذیر را هدف قرار می‌دهند، مانند افراد مسن، که به راحتی فریب کمپین‌های فیشینگ می‌خورند. هوش مصنوعی با تحلیل داده‌های شخصی و پروفایل‌های رسانه‌های اجتماعی، حملات هدفمند و دقیق ایجاد می‌کند که نگرانی‌های اخلاقی درباره حفاظت از این گروه‌ها و جلوگیری از سوءاستفاده‌ها را به وجود می‌آورد. [۵۴,۵۵]
- **تسلیح مدل‌های هوش مصنوعی:** با دسترسی بیشتر به فناوری‌های هوش مصنوعی، نگرانی‌ها در مورد استفاده مخرب از آن‌ها افزایش یافته است. مدل‌های متن‌باز مانند فناوری دیپ‌فیک می‌توانند برای ساخت ویدیوهای جعلی، گسترش اطلاعات غلط یا ایجاد بی‌ثباتی سیاسی به کار روند، و این دسترسی آسان نگرانی‌هایی درباره تسلیح هوش مصنوعی توسط بازیگران مختلف ایجاد می‌کند. [۴۲]
- **تعصب در حملات هوش مصنوعی:** نگرانی اخلاقی دیگر، پتانسیل تعصبات در سیستم‌های هوش مصنوعی است که از داده‌های ناقص یا غیرنمایانگر آموزش دیده‌اند. این تعصبات می‌توانند به فرآیندهای تصمیم‌گیری ناعادلانه منجر شوند، مانند هدف‌گیری گروه‌های خاص در حملات سایبری بر اساس ویژگی‌های جمعیتی، که نابرابری‌های اجتماعی را تشدید می‌کند. [58,43]

این نگرانی‌های اخلاقی بر لزوم توسعه، استقرار و حکمرانی مسئولانه فناوری‌های هوش مصنوعی تأکید دارند. با ترویج شفافیت، مسئولیت‌پذیری و عدالت در سیستم‌های هوش مصنوعی، می‌توانیم خطرات ناشی از استفاده‌های مخرب هوش مصنوعی را کاهش دهیم و اطمینان حاصل کنیم که هوش مصنوعی به‌جای آسیب رساندن به جامعه، به نفع آن عمل می‌کند. [۴۵]

### خلاءهای قانونی

پیشرفت سریع فناوری‌های هوش مصنوعی، به‌ویژه در حوزه امنیت سایبری، از توسعه چارچوب‌های قانونی و نظارتی که بتوانند به‌طور مؤثر استفاده از آن‌ها را تنظیم کنند، پیشی گرفته است. فاصله‌های قانونی در برخورد با حملات سایبری مبتنی بر هوش مصنوعی چالشی بزرگ برای قانون‌گذاران و حرفه‌ای‌های امنیت سایبری ایجاد کرده است. چندین مسئله کلیدی باید برای پر کردن این فاصله‌ها مورد توجه قرار گیرد:

- **کمبود سیاست‌های جامع:** قوانین فعلی امنیت سایبری عمدتاً بر تهدیدات سنتی مانند ویروس‌ها و بدافزارها تمرکز دارند و به پیچیدگی‌های جدید ناشی از هوش مصنوعی توجه نمی‌کنند. حملات مبتنی بر هوش مصنوعی مانند شبیه‌سازی خودکار، بدافزارهای خودتکثیر شونده و دیپ‌فیک‌ها در قوانین فعلی پوشش داده نمی‌شوند. علاوه بر این، نگرانی‌های اخلاقی مربوط به شفافیت، عدالت و پاسخگویی سیستم‌های هوش مصنوعی نادیده گرفته می‌شود. [60,40]

- **چالش‌های صلاحیت قضائی:** حملات سایبری اغلب به‌صورت فرامرزی انجام می‌شوند، و مهاجمان از مناطق با نظارت قانونی ضعیف سوءاستفاده می‌کنند. این امر شناسایی و تعقیب قانونی آن‌ها را برای آژانس‌های اجرایی پیچیده می‌کند. در حملات مبتنی بر هوش مصنوعی، باتنت‌ها معمولاً از راه دور و از کشورهای دارای قوانین امنیت سایبری ناکارآمد کنترل می‌شوند، که نیاز به همکاری بین‌المللی را برای مقابله با این تهدیدات دوچندان می‌کند. [۱۱]

- **پاسخگویی و مسئولیت:** تعیین مسئولیت در حملات سایبری مبتنی بر هوش مصنوعی پیچیده است به دلیل خودمختاری سیستم‌های هوش مصنوعی. این سیستم‌ها می‌توانند به‌طور مستقل تصمیم‌گیری کنند، که باعث ابهام در تعیین مسئولیت در صورت بروز آسیب می‌شود. برای مثال، در حملات DDoS با استفاده از هوش مصنوعی، مشخص نیست که آیا توسعه‌دهنده، کاربر یا خود سیستم باید مسئول شناخته شوند. این ابهام مانع از تضمین عدالت و جلوگیری از سوءاستفاده از فناوری می‌شود. [۱۸]

- **فرایندهای قانونی کند:** تکامل سریع فناوری هوش مصنوعی از توانایی دولت‌ها برای ایجاد و اجرای مقررات مرتبط پیشی گرفته است. قوانین فعلی ممکن است به سرعت منسوخ شوند، زیرا مهاجمین از فاصله‌های قانونی بهره‌برداری می‌کنند. بنابراین، سیاست‌گذاران باید چارچوب‌های قانونی چابکی ایجاد کنند که بتوانند با پیشرفت‌های تکنولوژیکی همگام شوند و با تهدیدات جدید مقابله کنند. [۳۷]

- **چالش‌های همکاری بین‌المللی:** تنظیم مقررات مؤثر برای حملات سایبری مبتنی بر هوش مصنوعی نیاز به همکاری بین‌المللی دارد، اما تفاوت‌ها در قوانین، منافع ملی و اولویت‌های فناوری مانع از ایجاد استانداردهای جهانی هماهنگ می‌شود. هنوز چارچوب بین‌المللی یکپارچه‌ای برای مقابله با این تهدیدات وجود ندارد. [47,36]

برای مقابله با این فاصله‌های قانونی، ضروری است که نهادهای بین‌المللی، مانند سازمان ملل و اتحادیه اروپا، با بخش خصوصی، دانشگاه‌ها و سازمان‌های امنیت سایبری همکاری نزدیکی داشته باشند تا چارچوب‌های قانونی جامع و انعطاف‌پذیری توسعه دهند که بتوانند به‌طور مؤثر به چالش‌هایی که هوش مصنوعی در حملات سایبری ایجاد می‌کند، رسیدگی کنند. [۳۷]

## مدل های پیش بینی برای جلوگیری از حملات

### استفاده از هوش مصنوعی برای دفاع

با پیچیده‌تر و گسترده‌تر شدن حملات سایبری، هوش مصنوعی (AI) و یادگیری ماشین (ML) ابزارهای ارزشمندی برای تقویت قابلیت‌های دفاعی سازمان‌ها فراهم می‌کنند. ادغام هوش مصنوعی در ابزارهای امنیت سایبری هدف دارد تا سیستم‌های هوشمندی بسازد که بتوانند حملات را در زمان واقعی پیش‌بینی، شناسایی و پاسخ دهند و به این ترتیب اشتباهات انسانی را به حداقل رسانده، زمان‌های پاسخگویی را کاهش داده و تاب‌آوری کلی سیستم‌ها را بهبود بخشند. چند روش اضافی که هوش مصنوعی به دفاع کمک می‌کند عبارتند از: [۱۱]

- **حفاظت از End Point ها با هوش مصنوعی:** راه‌حل‌های امنیتی نقطه پایانی مبتنی بر هوش مصنوعی برای محافظت از دستگاه‌ها در برابر تهدیدات مخرب استفاده می‌شوند. این سیستم‌ها از یادگیری ماشین برای شناسایی و مسدود کردن تهدیدات شناخته‌شده و ناشناخته استفاده می‌کنند و به‌طور خودکار واکنش نشان می‌دهند. از جمله قابلیت‌های این سیستم‌ها، شناسایی رفتارهای غیرمعمول و ارائه بینش برای اقدام پیشگیرانه است. [۴۶]

- **یادگیری عمیق برای تحلیل بدافزار:** سیستم‌های مبتنی بر هوش مصنوعی با استفاده از یادگیری عمیق و شبکه‌های عصبی کانولوشنی (CNN) در شناسایی بدافزارها موفقیت‌هایی داشته‌اند. این سیستم‌ها قادر به تحلیل رفتار فایل‌ها، ترافیک شبکه و فراخوانی‌های سیستم برای شناسایی بدافزارهای جدید هستند و به‌طور مداوم با داده‌های جدید بهبود می‌یابند، بدون نیاز به دخالت انسان. [۲]

- **استفاده از هوش مصنوعی برای شناسایی تهدیدات RealTime در محیط‌های ابری:** رایانش ابری چالش‌های امنیتی جدیدی ایجاد کرده است. سیستم‌های امنیتی سنتی در محیط‌های ابری کارآمد نیستند، اما سیستم‌های هوش مصنوعی می‌توانند با نظارت مستمر، شناسایی و کاهش تهدیدات در زمان واقعی را انجام دهند. این سیستم‌ها با استفاده از الگوریتم‌های داده‌کاوی، تهدیدات غیرعادی مانند تلاش‌های ورودی غیرمجاز یا حرکت مشکوک داده‌ها را شناسایی کرده و اقدامات مقابله‌ای مانند ایزوله کردن ماشین‌های مجازی را به‌طور خودکار انجام می‌دهند. [2,4,5]

- **چت‌بات‌های امنیت سایبری برای جمع‌آوری اطلاعات تهدید و گزارش حوادث:** چت‌بات‌های مبتنی بر هوش مصنوعی در امنیت سایبری می‌توانند اطلاعات تهدید را جمع‌آوری و تحلیل کنند، ارزیابی ریسک انجام دهند و راهنمایی‌هایی برای پاسخ به حوادث ارائه دهند. این چت‌بات‌ها همچنین ارتباطات بین تیم‌های امنیتی و ذینفعان را تسهیل کرده و با استفاده از یادگیری ماشین، پاسخ‌های خود را بهبود می‌دهند. به‌عنوان مثال، در مواجهه با نفوذ داده، چت‌بات می‌تواند کارکنان را در مهار نفوذ و ایمن‌سازی حساب‌ها هدایت کند. [2,6,5]

### مدل‌های پیش‌بینی

استفاده از مدل‌های پیش‌بینی در امنیت سایبری با افزایش حجم، پیچیدگی و سرعت حملات سایبری اهمیت بیشتری پیدا کرده است. مدل‌های پیش‌بینی به سازمان‌ها این امکان را می‌دهند که تهدیدات بالقوه را شناسایی کرده، عوامل ریسک را تحلیل کنند و اقدامات پیشگیرانه را پیش از آن که حمله باعث آسیب قابل توجهی شود، انجام دهند. این مدل‌ها داده‌های تاریخی، یادگیری ماشین و روش‌های آماری را ترکیب کرده تا احتمال وقوع سناریوهای مختلف حمله را پیش‌بینی کنند. در اینجا برخی از روش‌هایی که مدل‌های پیش‌بینی به امنیت سایبری کمک می‌کنند آورده شده است:[1,2,3]

- **پیش‌بینی نیت حملات سایبری بر اساس رویدادهای ژئوپولیتیکی:** مدل‌های پیش‌بینی مبتنی بر هوش مصنوعی قادرند احتمال حملات سایبری را بر اساس رویدادهای جهانی ژئوپولیتیکی ارزیابی کنند. این مدل‌ها می‌توانند با تحلیل داده‌های تاریخی حملات، هشدارهای زودهنگام برای سازمان‌ها ارسال کرده و به آژانس‌های دولتی کمک کنند تا احتمال حملات سایبری به زیرساخت‌های حیاتی یا دارایی‌های امنیت ملی را پیش‌بینی کنند. [2,4,5]

- **ارزیابی ریسک برای فروشندگان و زنجیره تأمین‌های ثالث:** مدل‌های پیش‌بینی مبتنی بر هوش مصنوعی برای ارزیابی ریسک امنیتی فروشندگان و شرکای ثالث به‌ویژه در صنایع حساس استفاده می‌شوند. این سیستم‌ها با تحلیل داده‌های حملات سایبری و آسیب‌پذیری‌های زنجیره تأمین، می‌توانند ریسک‌های بالقوه مربوط به فروشندگان را پیش‌بینی کرده و به سازمان‌ها کمک کنند تا تصمیمات آگاهانه‌تری درباره همکاری با آن‌ها بگیرند. این مدل‌ها می‌توانند ارزیابی کنند که آیا یک تأمین‌کننده دچار نفوذ داده شده یا از شیوه‌های امنیتی مناسب پیروی می‌کند. [2,4,5]

- **پیش‌بینی حملات در زمان واقعی با سیستم‌های SIEM تقویت‌شده توسط هوش مصنوعی:** سیستم‌های SIEM تقویت‌شده با هوش مصنوعی از تحلیل پیش‌بینی برای شناسایی تهدیدات سایبری در زمان واقعی استفاده می‌کنند. این سیستم‌ها از یادگیری ماشین، پردازش زبان طبیعی، تحلیل سری‌های زمانی و تحلیل خوشه‌ای برای شناسایی الگوهای حمله و ناهنجاری‌ها بهره می‌برند. با تشخیص الگوهای حملات در حال وقوع، این سیستم‌ها می‌توانند هشدارهای عملی ارائه دهند و به تیم‌های امنیتی کمک کنند تا اقدامات پیشگیرانه مانند مسدود کردن IPهای مشکوک یا به‌روزرسانی فایروال‌ها را انجام دهند. [۸]

## حل مدل و ارائه مثال نمونه

### استفاده از هوش مصنوعی برای پیش‌بینی حملات باج‌افزار

حملات باج‌افزار به یکی از رایج‌ترین انواع جرایم سایبری تبدیل شده‌اند. مدل‌های پیش‌بینی می‌توانند به شناسایی احتمال وقوع حمله باج‌افزار قبل از وقوع آن کمک کنند، از طریق تحلیل عواملی مانند وجود پیوست‌های ایمیل مخرب، فعالیت غیرمعمول شبکه، و داده‌های تاریخی در مورد انواع شناخته‌شده باج‌افزارها. سیستم‌های هوش مصنوعی می‌توانند از الگوریتم‌های یادگیری ماشین برای شناسایی و علامت‌گذاری این علائم هشداردهنده اولیه استفاده کنند، به تیم‌های امنیتی این فرصت را می‌دهند که حمله را قبل از گسترش متوقف کنند. در برخی موارد، مدل‌های پیش‌بینی همچنین می‌توانند شدت حمله و آسیب بالقوه آن را ارزیابی کنند، که به سازمان‌ها این امکان را می‌دهد تا پاسخ خود را اولویت‌بندی کنند.

با استفاده از مدل‌های پیش‌بینی، سازمان‌ها می‌توانند یک گام از هکرها جلوتر باشند و ریسک‌ها را پیش از آنکه منجر به عواقب جدی شوند، شناسایی و کاهش دهند. [۱۵]

- **همکاری بین دولت‌ها و سازمان‌ها:** همکاری بین دولت‌ها، سازمان‌های بین‌المللی و نهادهای بخش خصوصی برای ایجاد یک دفاع جمعی در برابر تهدیدات سایبری ضروری است. امنیت سایبری ذاتاً یک مسئله جهانی است که نیازمند تلاش‌های هماهنگ برای مقابله با ماهیت پیچیده و مداوم حملات سایبری است. در ادامه بحث قبلی، روش‌های بیشتری که همکاری می‌تواند امنیت سایبری را تقویت کند، آورده شده است: [19]

- **شبکه‌های جهانی تبادل اطلاعات:** دولت‌ها و سازمان‌های خصوصی می‌توانند در شبکه‌های جهانی تبادل اطلاعات امنیت سایبری شرکت کنند که به اشتراک‌گذاری جزئیات حملات، روش‌ها، تاکتیک‌ها و آسیب‌پذیری‌ها را تسهیل می‌کند. این همکاری‌ها اطلاعات تهدید جامع‌تری فراهم می‌کنند که به سازمان‌ها کمک می‌کند از هکرها پیشی بگیرند. نمونه‌هایی از این همکاری‌ها شامل مجمع جهانی تخصص در زمینه سایبری (GFCE) و آژانس امنیت سایبری اتحادیه اروپا (ENISA) است که برای بهبود دفاع در برابر تهدیدات نوظهور و ارتقاء بهترین شیوه‌ها در حال تلاش هستند. [10,23]

- **همکاری‌های تحقیق و توسعه در زمینه امنیت سایبری:** همکاری بین دولت‌ها و سازمان‌ها می‌تواند شامل تلاش‌های مشترک تحقیق و توسعه (R&D) برای ایجاد فناوری‌های جدید تقویت‌کننده دفاع در برابر حملات سایبری مبتنی بر هوش مصنوعی باشد. با همکاری مؤسسات آکادمیک، نهادهای دولتی و صنعت خصوصی، می‌توان نوآوری‌هایی در زمینه الگوریتم‌های یادگیری ماشین برای بهبود تشخیص تهدیدات ایجاد کرد. این رویکرد باعث توسعه فناوری‌های پیشرفته امنیت سایبری می‌شود که برای جامعه مفید است و به ایجاد چارچوب‌های امنیتی مشترک و ابزارهای هماهنگ در صنایع مختلف کمک می‌کند. [10,23]

- **دیپلماسی سایبری برای همکاری‌های بین‌المللی در زمینه امنیت سایبری:** از آنجا که حملات سایبری مرزهای ملی را می‌شکافند، دیپلماسی سایبری برای ایجاد چارچوب‌هایی جهت همکاری بین کشورها در زمینه امنیت سایبری ضروری است. این همکاری شامل مذاکره در مورد توافق‌نامه‌های بین‌المللی، تبادل اطلاعات تهدید و هماهنگی پاسخ‌ها به حوادث سایبری است. از طریق این توافق‌ها، کشورها می‌توانند پروتکل‌هایی برای پاسخ به حملات سایبری مرزی، نسبت دادن مسئولیت و

مقابله با تهدیدات سایبری حمایت‌شده از دولت‌ها ایجاد کنند. این همکاری دیپلماتیک به ایجاد پاسخ جهانی مؤثر به حملات سایبری مبتنی بر هوش مصنوعی کمک خواهد کرد. [۸،۱۳]

این تلاش‌های جمعی در زمینه امنیت سایبری به دولت‌ها و سازمان‌ها این امکان را می‌دهند تا تهدید رو به رشد حملات سایبری مبتنی بر هوش مصنوعی را به‌طور مؤثرتری مقابله کنند و یک اکوسیستم جهانی امنیت سایبری ایجاد کنند که برای تمامی ذینفعان سودمند باشد. [13,8]

## نمونه‌هایی از مدل‌های پیش‌بینی و مکانیزم‌های دفاعی مبتنی بر هوش مصنوعی

مدل‌های پیش‌بینی و مکانیزم‌های دفاعی مبتنی بر هوش مصنوعی (AI) در حوزه امنیت سایبری نقش برجسته‌ای دارند. این سیستم‌ها از الگوریتم‌ها و مدل‌های پیچیده برای شبیه‌سازی تهدیدات، پیش‌بینی حملات و طراحی راهکارهای مقابله‌ای استفاده می‌کنند. در اینجا به برخی از این مدل‌ها و مکانیزم‌ها اشاره می‌کنیم:

- **مدل‌های پیش‌بینی مبتنی بر یادگیری ماشین:** مدل‌های یادگیری ماشین به‌ویژه الگوریتم‌های نظارت‌شده و غیرنظارت‌شده، برای شناسایی الگوهای مشکوک و پیش‌بینی تهدیدات سایبری به کار می‌روند. یکی از نمونه‌های معروف، مدل‌های پیش‌بینی مبتنی بر دسته‌بندی هستند که می‌توانند حملات مختلف از جمله حملات DDoS، ویروس‌ها و حملات فیشینگ را شناسایی کنند.

- **الگوریتم‌های دسته‌بندی:** الگوریتم‌هایی مانند درخت تصمیم، جنگل تصادفی و پشتیبانی بردار ماشین (SVM) برای طبقه‌بندی فعالیت‌های شبکه به انواع مختلف از جمله تهدیدات یا رفتارهای طبیعی استفاده می‌شوند.

- **مدل‌های شبیه‌سازی تهدید:** این مدل‌ها از داده‌های تاریخی برای شبیه‌سازی حملات مختلف و پیش‌بینی تهدیدات استفاده می‌کنند. به عنوان مثال، استفاده از داده‌های وقایع امنیتی و شبکه برای شبیه‌سازی حملات پیشرفته می‌تواند به شناسایی و پیش‌بینی حملات جدید کمک کند.

- **سیستم‌های تشخیص نفوذ (IDS) مبتنی بر هوش مصنوعی:** این سیستم‌ها با استفاده از الگوریتم‌های یادگیری ماشین و هوش مصنوعی به شناسایی و پیش‌بینی حملات در شبکه‌ها پرداخته و از آن‌ها دفاع می‌کنند. در این سیستم‌ها، به جای استفاده از الگوهای ثابت برای شناسایی حملات، سیستم‌های هوش مصنوعی می‌توانند تهدیدات جدید و ناشناخته را شبیه‌سازی و شناسایی کنند.

- **شبکه‌های عصبی مصنوعی (ANN):** شبکه‌های عصبی به‌ویژه مدل‌های یادگیری عمیق، می‌توانند به عنوان یک سیستم تشخیص نفوذ برای شناسایی الگوهای پیچیده حملات سایبری و بدافزارها استفاده شوند.

- **سیستم‌های تشخیص مبتنی بر رفتار:** این سیستم‌ها فعالیت‌های معمول شبکه را یاد می‌گیرند و در صورتی که یک رفتار غیرعادی مشاهده شود، به‌عنوان یک تهدید تلقی می‌شود.

- **سیستم‌های دفاعی مبتنی بر هوش مصنوعی برای مقابله با حملات DDoS:** حملات توزیع‌شده انکار خدمات (DDoS) یکی از رایج‌ترین تهدیدات در دنیای سایبری هستند که می‌توانند خدمات آنلاین را مختل کنند. استفاده از هوش مصنوعی برای پیش‌بینی و جلوگیری از چنین حملاتی به‌طور فزاینده‌ای رایج شده است.

- **الگوریتم‌های پیش‌بینی و شبیه‌سازی ترافیک:** با استفاده از داده‌های تاریخی ترافیک شبکه، الگوریتم‌های هوش مصنوعی می‌توانند ترافیک طبیعی را از ترافیک حملات DDoS تمیز دهند و در زمان واقعی آن‌ها را شناسایی کنند.

- **مدیریت خودکار منابع:** سیستم‌های هوش مصنوعی قادر به شبیه‌سازی و پیش‌بینی حملات DDoS هستند و می‌توانند منابع شبکه را به‌صورت خودکار مدیریت کنند تا حملات را کاهش دهند.

- **مکانیزم‌های دفاعی مبتنی بر تحلیل تهدید و ارزیابی ریسک:** مدل‌های هوش مصنوعی می‌توانند برای تحلیل تهدیدات، ارزیابی ریسک‌ها و پیش‌بینی آسیب‌های احتمالی ناشی از حملات سایبری استفاده شوند. این سیستم‌ها می‌توانند نقاط ضعف موجود در زیرساخت‌های شبکه را شناسایی کرده و اقدامات پیشگیرانه را پیشنهاد دهند.

- **مدل‌های ارزیابی ریسک مبتنی بر یادگیری ماشین:** این مدل‌ها با تحلیل داده‌های تهدیدات گذشته و ارزیابی آسیب‌ها، می‌توانند به سازمان‌ها در تصمیم‌گیری‌های امنیتی کمک کنند. الگوریتم‌هایی مانند یادگیری عمیق و تحلیل خوشه‌ای برای شبیه‌سازی تهدیدات و ارزیابی ریسک‌ها به‌کار می‌روند.

- **دفاع‌های هوش مصنوعی در برابر حملات فیشینگ:** حملات فیشینگ از روش‌های رایج برای سرقت اطلاعات حساس هستند. سیستم‌های هوش مصنوعی می‌توانند به شناسایی و مقابله با این نوع حملات کمک کنند.

- **مدل‌های شناسایی ایمیل‌های فیشینگ:** با استفاده از الگوریتم‌های پردازش زبان طبیعی (NLP) و یادگیری ماشین، سیستم‌های هوش مصنوعی قادر به شناسایی ایمیل‌ها و پیغام‌های مشکوک هستند. این مدل‌ها می‌توانند متنی که به نظر فیشینگ می‌آید را تحلیل کنند و کاربران را از خطرات احتمالی آگاه سازند.

این نمونه‌ها تنها گوشه‌ای از کاربردهای هوش مصنوعی در مقابله با تهدیدات سایبری است و همچنان در حال توسعه هستند. هوش مصنوعی با قابلیت شناسایی الگوهای پیچیده و پیش‌بینی تهدیدات نوین، ابزاری قدرتمند برای ایجاد سیستم‌های دفاعی قوی‌تر در برابر تهدیدات پیچیده‌تر است.

## یافته ها

هوش مصنوعی (AI) در دنیای امروز به‌عنوان یک ابزار کلیدی در پیشگیری و مقابله با حملات سایبری شناخته می‌شود. توانایی‌های پیشرفته‌ی AI در تحلیل داده‌ها، شناسایی الگوهای پیچیده و تصمیم‌گیری سریع به‌طور فزاینده‌ای برای حفاظت از سیستم‌های اطلاعاتی در برابر تهدیدات سایبری مورد استفاده قرار می‌گیرد. در این راستا، نقش هوش مصنوعی در پیشگیری و مقابله با حملات سایبری را می‌توان در چند جنبه مهم تحلیل کرد:

- **شناسایی تهدیدات و حملات جدید:** یکی از بزرگ‌ترین چالش‌ها در امنیت سایبری، شناسایی حملات جدید و ناشناخته است. بسیاری از حملات سایبری پیچیده و با استفاده از تکنیک‌های جدید طراحی می‌شوند که می‌تواند از سیستم‌های امنیتی قدیمی عبور کند. در این زمینه، هوش مصنوعی با استفاده از مدل‌های یادگیری ماشین و یادگیری عمیق، قادر است الگوهای مشکوک و حملات ناشناخته را شبیه‌سازی و شناسایی کند.
- **سیستم‌های تشخیص نفوذ (IDS):** هوش مصنوعی با شبیه‌سازی رفتار طبیعی شبکه، قادر است حملات جدید را که تاکنون شناسایی نشده‌اند، شناسایی کند.
- **مدل‌های پیش‌بینی:** با استفاده از داده‌های تاریخی و مدل‌های هوش مصنوعی، می‌توان حملات آینده را پیش‌بینی کرده و اقدامات پیشگیرانه انجام داد.
- **سیستم‌های دفاعی همکاری‌پذیر:** سیستم‌های هوش مصنوعی می‌توانند در شبکه‌های سازمان‌ها همکاری کنند تا داده‌های تهدید را تبادل کنند و تدابیر امنیتی را به‌طور بهینه تنظیم کنند. این همکاری می‌تولند فرصت‌هایی برای اشتراک‌گذاری اطلاعات تهدید و استراتژی‌های دفاعی بین شرکت‌ها ایجاد کند. [۶]
- **محاسبات کوانتومی و هوش مصنوعی:** ادغام محاسبات کوانتومی با هوش مصنوعی می‌تواند به‌طور بنیادینی دفاع‌های امنیت سایبری را تغییر دهد. با قدرت پردازشی بالا، محاسبات کوانتومی می‌تواند تحلیل داده‌های بزرگ را سریع‌تر انجام دهد و در ترکیب با الگوریتم‌های هوش مصنوعی، شناسایی تهدیدات و بهینه‌سازی استراتژی‌های دفاعی را در زمان واقعی فراهم کند. [۱۱]
- **هماهنگی امنیتی خودکار:** اتوماسیون عملیات امنیتی با هوش مصنوعی اهمیت فزاینده‌ای یافته است. سیستم‌های پاسخ‌دهی خودکار به حوادث از یادگیری ماشین برای تحلیل و واکنش به تهدیدات بدون دخالت انسان استفاده می‌کنند. راه‌حل‌های پیشرفته امنیتی مبتنی بر هوش مصنوعی بر هماهنگی ابزارها و فرآیندها تمرکز دارند تا شبکه‌ای دفاعی، یکپارچه و خودکار ایجاد کنند که سریع‌تر از تحلیل‌گران انسانی عمل کند. [۲۳]

**جدول ۱:انواع روش های هوش مصنوعی در حملات**

| جنبه ها | توضیحات |
|---|---|

| | |
|---|---|
| **شناسایی تهدیدات جدید** | شناسایی حملات ناشناخته با استفاده از یادگیری ماشین و یادگیری عمیق. |
| **سیستم‌های تشخیص نفوذ (IDS)** | شبیه‌سازی رفتار طبیعی شبکه و شناسایی حملات جدید و ناشناخته. |
| **مدل‌های پیش‌بینی** | پیش‌بینی حملات آینده با استفاده از داده‌های تاریخی و الگوریتم‌های هوش مصنوعی. |
| **سیستم‌های دفاعی همکاری‌پذیر** | اشتراک‌گذاری اطلاعات تهدید بین سازمان‌ها و تنظیم تدابیر امنیتی بهینه. |
| **محاسبات کوانتومی و هوش مصنوعی** | استفاده از قدرت پردازش کوانتومی برای تحلیل سریع‌تر داده‌های بزرگ. |
| **هماهنگی امنیتی خودکار** | تحلیل و واکنش خودکار به تهدیدات بدون نیاز به مداخله انسانی. |

### هوش مصنوعی برای تحلیل اطلاعات تهدید سایبری

اطلاعات تهدید سایبری (CTI) شامل جمع‌آوری، ارزیابی و انتشار داده‌ها در مورد تهدیدات سایبری بالقوه و فعال است. این فرآیند به‌طور سنتی دستی بوده و از متخصصان امنیتی خواسته می‌شود که حجم زیادی از داده‌ها را از منابع مختلف مانند ترافیک شبکه، رسانه‌های اجتماعی و نظارت بر وب تاریک بررسی کنند تا تهدیدات نوظهور را شناسایی کنند. هوش مصنوعی پتانسیل اتوماسیون بسیاری از این فرآیندها را دارد، که آن را کارآمدتر و دقیق‌تر می‌کند. [۳۶]

- **پردازش زبان طبیعی (NLP) و استخراج متن:** هوش مصنوعی با استفاده از NLP و استخراج متن، می‌تواند از منابع عمومی و آنلاین اطلاعات تهدید سایبری را استخراج کرده و با تحلیل احساسات، تهدیدات جدید را پیش از شناخته‌شدن شناسایی کند. [۴۴]
- **تحلیل رفتاری:** هوش مصنوعی با تحلیل داده‌های تاریخی حملات و رفتار مهاجمان، می‌تواند حرکات آینده آن‌ها را پیش‌بینی کرده و استراتژی‌های دفاعی پیشگیرانه ارائه دهد. [۴۶]
- **همبستگی خودکار تهدیدات:** هوش مصنوعی با پردازش سریع داده‌ها، ارتباطات میان تهدیدات ظاهراً نامرتبط را شناسایی کرده و بینش‌هایی جامع درباره حملات پیشرفته (APTs) ارائه می‌دهد که به کنترل آن‌ها پیش از گسترش کمک می‌کند. [۱۵]
- **پیش‌بینی تهدیدات:** هوش مصنوعی با تحلیل داده‌های تاریخی و الگوهای حمله، حملات سایبری احتمالی را پیش‌بینی کرده و به سازمان‌ها امکان می‌دهد اقدامات پیشگیرانه برای کاهش خطرات و تقویت دفاع‌های خود انجام دهند. [۲۶]

### مقایسه نتایج با پژوهش‌های پیشین

مقایسه نتایج یک پژوهش با نتایج پژوهش‌های پیشین، به‌ویژه در زمینه‌ای مانند نقش هوش مصنوعی در پیشگیری و مقابله با حملات سایبری، برای تحلیل روندهای پیشرفت تکنولوژی و ارزیابی قابلیت‌های جدید بسیار حائز اهمیت است. در اینجا می‌توانیم چند جنبه مختلف مقایسه را بررسی کنیم:

### پیشرفت در دقت و کارآیی مدل‌ها:

در پژوهش‌های پیشین، شناسایی تهدیدات و حملات سایبری با استفاده از تکنیک‌های سنتی مانند الگوریتم‌های مبتنی بر قوانین یا تحلیل‌های آماری انجام می‌شد. در مقایسه با این روش‌ها، پژوهش‌های جدید که به کارگیری یادگیری ماشین و یادگیری عمیق را شامل می‌شوند، دقت بیشتری در شناسایی تهدیدات ناشناخته و پیچیده دارند.

- **پژوهش‌های پیشین:** استفاده از الگوریتم‌های ساده‌تر و کمتر پیشرفته، معمولاً محدود به شناسایی تهدیدات شناخته‌شده بوده است.
- **پژوهش‌های جدید:** الگوریتم‌های هوش مصنوعی قادرند تهدیدات ناشناخته را شبیه‌سازی کنند و مدل‌های یادگیری عمیق به‌طور مؤثری الگوهای پیچیده را شناسایی نمایند.

### توسعه ابزارهای خودکار و واکنش‌های آنی:

یکی از تفاوت‌های عمده بین پژوهش‌های قدیمی و جدید در این حوزه، در استفاده از سیستم‌های خودکار پاسخ‌دهی به تهدیدات است. در گذشته، سیستم‌ها معمولاً نیاز به مداخله انسانی برای تشخیص و پاسخ به تهدیدات داشتند.

- **پژوهش‌های پیشین:** واکنش‌ها به تهدیدات بیشتر وابسته به مداخلات دستی و تحلیل‌های دستی بود.
- **پژوهش‌های جدید:** سیستم‌های مبتنی بر هوش مصنوعی می‌توانند به‌طور خودکار اقدامات دفاعی انجام دهند و حتی برخی از آن‌ها از سیستم‌های خودترمیمی بهره می‌برند که به‌طور خودکار آسیب‌های وارد شده را اصلاح می‌کنند.

### مقاومت در برابر حملات پیچیده‌تر:

با توجه به پیشرفت تکنولوژی‌های مورد استفاده در حملات سایبری، حملات امروزی به‌طور فزاینده‌ای پیچیده‌تر و پیشرفته‌تر شده‌اند. پژوهش‌های جدید نشان می‌دهند که استفاده از هوش مصنوعی می‌تواند این پیچیدگی‌ها را مدیریت کند و در مقایسه با روش‌های قدیمی‌تر، کارآمدتر باشد.

- **پژوهش‌های پیشین:** حملات سایبری اغلب با استفاده از ابزارهای ساده‌تر و با هدف دستیابی به اطلاعات عمومی صورت می‌گرفت.

- **پژوهش‌های جدید:** با ظهور حملات پیچیده مانند APT (Advanced Persistent Threats)، سیستم‌های هوش مصنوعی قادرند به شبیه‌سازی و پیش‌بینی این نوع تهدیدات بپردازند و آن‌ها را در مراحل اولیه شناسایی کنند.

**استفاده از پردازش زبان طبیعی (NLP) در شناسایی حملات اجتماعی:**

در گذشته، حملات فیشینگ و حملات اجتماعی عمدتاً به‌صورت دستی شناسایی می‌شدند و الگوریتم‌های هوش مصنوعی در این زمینه کمتر به کار می‌رفتند. پژوهش‌های جدید نشان داده‌اند که با استفاده از پردازش زبان طبیعی (NLP)، می‌توان حملات فیشینگ و حملات اجتماعی را با دقت بیشتری شناسایی کرد.

- **پژوهش‌های پیشین:** شناسایی حملات اجتماعی به‌صورت دستی و با استفاده از قواعد از پیش تعریف‌شده انجام می‌شد.
- **پژوهش‌های جدید:** با استفاده از الگوریتم‌های NLP و تحلیل محتوای پیام‌ها و ایمیل‌ها، حملات فیشینگ به‌طور خودکار شناسایی می‌شوند.

**پیش‌بینی و ارزیابی ریسک:**

پژوهش‌های پیشین در زمینه امنیت سایبری معمولاً روی شناسایی و واکنش به تهدیدات متمرکز بودند، اما در پژوهش‌های جدید، استفاده از هوش مصنوعی برای پیش‌بینی و ارزیابی ریسک‌های امنیتی به‌طور فزاینده‌ای مورد توجه قرار گرفته است.

- **پژوهش‌های پیشین:** بیشتر بر روی شناسایی تهدیدات موجود تمرکز داشتند و کمتر به ارزیابی ریسک و پیش‌بینی تهدیدات آینده پرداخته بودند.
- **پژوهش‌های جدید:** مدل‌های هوش مصنوعی قادرند ریسک‌های امنیتی را پیش‌بینی کنند و اقدامات پیشگیرانه برای مقابله با آن‌ها انجام دهند.

**مقایسه با سیستم‌های سنتی:**

استفاده از سیستم‌های سنتی در شناسایی حملات سایبری (مانند فایروال‌ها و سیستم‌های تشخیص نفوذ) به‌طور معمول محدود به شناسایی حملات شناخته‌شده بود. در مقایسه، سیستم‌های هوش مصنوعی جدید قادرند با شناسایی الگوهای ناشناخته و تحلیل رفتارهای غیرعادی، به‌طور پیشرفته‌تری به شناسایی و مقابله با تهدیدات بپردازند.

- **پژوهش‌های پیشین:** بیشتر به تحلیل و شناسایی حملات از طریق روش‌های قاعده‌محور (rule-based) می‌پرداختند.
- **پژوهش‌های جدید:** توانایی تحلیل رفتارهای غیرعادی و شناسایی حملات جدید و پیچیده‌تر، از ویژگی‌های کلیدی پژوهش‌های جدید به‌شمار می‌آید.

**جدول ۲:مقایسه پژوهش های قبلی با پژوهش های کارشده با هوش مصنوعی**

| ویژگی‌ها | پژوهش‌های پیشین | پژوهش‌های جدید |
|---|---|---|
| **دقت شناسایی تهدیدات** | محدود به تهدیدات شناخته‌شده | شناسایی تهدیدات ناشناخته و پیچیده با یادگیری عمیق. |
| **ابزارهای خودکار** | نیاز به مداخلات انسانی | پاسخ‌دهی و تحلیل خودکار با استفاده از AI. |
| **مقاومت در برابر حملات پیچیده** | ابزارهای ساده‌تر برای مقابله با حملات سنتی | مدیریت حملات پیشرفته مانند APT با شبیه‌سازی هوش مصنوعی. |
| **استفاده از NLP** | شناسایی دستی حملات اجتماعی | تحلیل خودکار پیام‌ها و ایمیل‌ها با NLP. |
| **پیش‌بینی ریسک** | تمرکز بر تهدیدات موجود | پیش‌بینی تهدیدات آینده و ارزیابی ریسک. |

## بحث و نتیجه‌گیری

هوش مصنوعی یک شمشیر دو لبه در دنیای امنیت سایبری است. در حالی که دوره جدیدی از قابلیت‌ها را برای مهاجمان سایبری معرفی کرده و امکان حملات سریع‌تر و پیچیده‌تر را فراهم کرده است، ابزارهای قدرتمندی نیز برای دفاع فراهم می‌آورد. آینده امنیت سایبری بستگی به یکپارچه‌سازی هوش مصنوعی در استراتژی‌های تهاجمی و دفاعی خواهد داشت، اما این امر باید با استفاده مسئولانه و اخلاقی از این فناوری‌ها همراه باشد. چارچوب‌های نظارتی باید به‌روزرسانی شوند تا تهدیدات رو به رشد ناشی از هوش مصنوعی را برطرف کنند و همکاری بین‌المللی برای مقابله مؤثر با تهدیدات سایبری مبتنی بر هوش مصنوعی ضروری است. تحقیقات آینده نقش کلیدی در پیشرفت توانمندی‌های هوش مصنوعی برای امنیت سایبری ایفا خواهد کرد و چالش‌های اخلاقی و فنی که به‌وجود می‌آید را نیز مورد بررسی قرار خواهد داد.

## منابع

1) National Academies of Sciences, Engineering, and Medicine, 2024. Large Language Models and Cybersecurity: Proceedings of a Workshop—in Brief.

2) Zhang, Z., Ning, H., Shi, F., Farha, F., Xu, Y., Xu, J., Zhang, F. and Choo, K.K.R., 2022. Artificial intelligence in cyber security: research advances, challenges, and opportunities. Artificial Intelligence Review, pp.1-25.

3) Babu, C.S., 2024. Adaptive AI for Dynamic Cybersecurity Systems: Enhancing Protection in a Rapidly Evolving Digital Landscap. In Principles and Applications of Adaptive Artificial Intelligence (pp. 52-72). IGI Global.

4) Kaur, R., Gabrijelčič, D. and Klobučar, T., 2023. Artificial intelligence for cybersecurity: Literature review and future research directions. Information Fusion, 97, p.101804.

5) Gupta, M., Akiri, C., Aryal, K., Parker, E. and Praharaj, L., 2023. From chatgpt to threatgpt: Impact of generative ai in cybersecurity and privacy. IEEE Access.

6) Wu, Y., Jiang, N., Pham, H.V., Lutellier, T., Davis, J., Tan, L., Babkin, P. and Shah, S., 2023, July. How effective are neural networks for fixing security vulnerabilities. In Proceedings of the 32nd ACM SIGSOFT International Symposium on Software Testing and Analysis (pp. 1282-1294).

7) Tushkanova, O., Levshun, D., Branitskiy, A., Fedorchenko, E., Novikova, E. and Kotenko, I., 2023. Detection of cyberattacks and anomalies in cyber-physical systems: Approaches, data sources, evaluation. Algorithms, 16(2), p.85.

8) Bouramdane, A.A., 2023. Cyberattacks in smart grids: challenges and solving the multi-criteria decision-making for cybersecurity options, including ones that incorporate artificial intelligence, using an analytical hierarchy process. Journal of Cybersecurity and Privacy, 3(4), pp.662-705.

9) Kaur, R., Gabrijelčič, D. and Klobučar, T., 2023. Artificial intelligence for cybersecurity: Literature review and future research directions. Information Fusion, 97, p.101804.

10) Adadi, A. and Berrada, M., 2018. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). IEEE access, 6, pp.52138-52160.

11) Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R. and Chatila, R., 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information fusion, 58, pp.82-115.

12) Bai, Y., Kadavath, S., Kundu, S., Askell, A., Kernion, J., Jones, A., Chen, A., Goldie, A., Mirhoseini, A., McKinnon, C. and Chen, C., 2022. Constitutional ai: Harmlessness from ai feedback. arXiv preprint arXiv:2212.08073.

13) Yan, K., Liu, X., Lu, Y. and Qin, F., 2022. A cyber-physical power system risk assessment model against cyberattacks. IEEE Systems Journal, 17(2), pp.2018-2028.

14) Chai, Y., Du, L., Qiu, J., Yin, L. and Tian, Z., 2022. Dynamic prototype network based on sample adaptation for few-shot malware detection. IEEE Transactions on Knowledge and Data Engineering, 35(5), pp.4754-4766.

15) Ng, M., Coopamootoo, K.P., Toreini, E., Aitken, M., Elliot, K. and van Moorsel, A., 2020, September. Simulating the effects of social presence on trust, privacy concerns & usage intentions in automated bots for finance. In 2020 IEEE European symposium on security and privacy workshops (EuroS&PW) (pp. 190-199). IEEE.

16) Bracken, B.K., Wolcott, J., Potoczny-Jones, I., Mosser, B.A., Griffith-Fillipo, I.R. and Areán, P.A., 2022. Detection and Remediation of Malicious Actors for Studies Involving Remote Data Collection. In HEALTHINF (pp. 377-383).

17) Bommasani, R., Hudson, D.A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M.S., Bohg, J., Bosselut, A., Brunskill, E. and Brynjolfsson, E., 2021. On the opportunities and risks of foundation models. arXiv preprint arXiv:2108.07258.

18) Cobbe, J., Veale, M. and Singh, J., 2023, June. Understanding accountability in algorithmic supply chains. In Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (pp. 1186-1197).

19) Nweke, L.O. and Yayilgan, S.Y., 2024. Opportunities and Challenges of Using Artificial Intelligence in Securing Cyber-Physical Systems. Artificial Intelligence for Security: Enhancing Protection in a Changing World, pp.131-164.

20) Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M. and Ghemawat, S., 2016. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467.

21) Karim, A., Shahroz, M., Mustofa, K., Belhaouari, S.B. and Joga, S.R.K., 2023. Phishing detection system through hybrid machine learning based on URL. IEEE Access, 11, pp.36805-36822.

22) Blanchard, P., El Mhamdi, E.M., Guerraoui, R. and Stainer, J., 2017. Machine learning with adversaries: Byzantine tolerant gradient descent. Advances in neural information processing systems, 30.

23) Fang, M., Cao, X., Jia, J. and Gong, N., 2020. Local model poisoning attacks to {Byzantine-Robust} federated learning. In 29th USENIX security symposium (USENIX Security 20) (pp. 1605-1622).

24) Wang, B., Yao, Y., Shan, S., Li, H., Viswanath, B., Zheng, H. and Zhao, B.Y., 2019, May. Neural cleanse: Identifying and mitigating backdoor attacks in neural networks. In 2019 IEEE symposium on security and privacy (SP) (pp. 707-723). IEEE.

25) Mienye, I.D. and Sun, Y., 2023. A deep learning ensemble with data resampling for credit card fraud detection. IEEE Access, 11, pp.30628-30638.

26) Falade, P.V., 2023. Decoding the threat landscape: Chatgpt, fraudgpt, and wormgpt in social engineering attacks. arXiv preprint arXiv:2310.05595.

27) Zhao, H., Zhou, W., Chen, D., Wei, T., Zhang, W. and Yu, N., 2021. Multi-attentional deepfake detection. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 2185-2194).

28) Distler, V., 2023, April. The influence of context on response to spear-phishing attacks: An in-situ deception study. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (pp. 1-18).

29) Tlili, A., Shehata, B., Adarkwah, M.A., Bozkurt, A., Hickey, D.T., Huang, R. and Agyemang, B., 2023. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart learning environments, 10(1), p.15.

30) Hutto, C. and Gilbert, E., 2014, May. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the international AAAI conference on web and social media (Vol. 8, No. 1, pp. 216-225).

31) Hutto, C. and Gilbert, E., 2014, May. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the international AAAI conference on web and social media (Vol. 8, No. 1, pp. 216-225).

32) Hutto, C. and Gilbert, E., 2014, May. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the international AAAI conference on web and social media (Vol. 8, No. 1, pp. 216-225).

33) Arslan, F., 2023. Deepfake Technology: A Criminological Literature Review. Sakarya Üniversitesi Hukuk Fakültesi Dergisi, 11(1), pp.701-720.

34) Geeng, C., 2020, April. Egregor: An eldritch privacy mental model for smart assistants. In Extended abstracts of the 2020 CHI conference on human factors in computing systems (pp. 1-9).

35) He, W., Xu, W., Ge, X., Han, Q.L., Du, W. and Qian, F., 2021. Secure control of multiagent systems against malicious attacks: A brief survey. IEEE Transactions on Industrial Informatics, 18(6), pp.3595-3608.

36) Guembe, B., Azeta, A., Misra, S., Osamor, V.C., Fernandez-Sanz, L. and Pospelova, V., 2022. The emerging threat of ai-driven cyber attacks: A review. Applied Artificial Intelligence, 36(1), p.2037254.

37) Elsisi, M., Altius, M., Su, S.F. and Su, C.L., 2023. Robust Kalman filter for position estimation of automated guided vehicles under cyberattacks. IEEE Transactions on Instrumentation and Measurement, 72, pp.1-12.

38) Derner, E. and Batistič, K., 2023. Beyond the safeguards: exploring the security risks of ChatGPT. arXiv preprint arXiv:2305.08005.

39) Guleria, A., Krishan, K., Sharma, V. and Kanchan, T., 2023. ChatGPT: ethical concerns and challenges in academics and research. The Journal of Infection in Developing Countries, 17(09), pp.1292-1299.

40) Wu, X., Duan, R. and Ni, J., 2024. Unveiling security, privacy, and ethical concerns of ChatGPT. Journal of Information and Intelligence, 2(2), pp.102-115.

41) Althobaiti, K., Wolters, M.K., Alsufyani, N. and Vaniea, K., 2023. Using clustering algorithms to automatically identify phishing campaigns. IEEE Access, 11, pp.96502-96513.

42) Johnson, B.D., 2017. The Weaponization of AI: A Glimpse into Future Threats. Computer, 50(10), pp.73-73.

43) Ferrara, E., 2024. The butterfly effect in artificial intelligence systems: Implications for AI bias and fairness. Machine Learning with Applications, 15, p.100525.

44) Patrikar, A.M., Mahenthiran, A. and Said, A., 2023, June. Leveraging synthetic data for AI bias mitigation. In Synthetic Data for Artificial Intelligence and Machine Learning: Tools, Techniques, and Applications (Vol. 12529, pp. 185-190). SPIE.

45) Trivedi, A., Kaur, E.K., Choudhary, C. and Barnwal, P., 2023, March. Should AI Technologies Replace the Human Jobs?. In 2023 2nd International Conference for Innovation in Technology (INOCON) (pp. 1-6). IEEE.

46) Veprytska, O. and Kharchenko, V., 2022, December. AI powered attacks against AI powered protection: Classification, scenarios and risk analysis. In 2022 12th International Conference on Dependable Systems, Services and Technologies (DESSERT) (pp. 1-7). IEEE.

# Artificial Intelligence in Cyber Attacks

Shahla maniei fard
Ph.D. Student of Computer Science, Artificial Intelligence, Azad University, Tehran

Abstract
The rapid advancement of artificial intelligence (AI) has had a significant impact onvarious domains, including cybersecurity. This study examines the theoretical dimensions of the use of AI in cyberattacks, highlighting its role in automating complex attack strategies, increasing the effectiveness of attacker methods, and implementing complex social engineering campaigns. Using machine learning algorithms, attackers can exploit vulnerabilities, bypass detection systems, and launch highly targeted and consistent attacks, posing a serious threat to digital infrastructure . However, the use of AI for malicious activities poses numerous ethical and technical challenges. These challenges include the difficulty in identifying AI-based threats, the ethical dilemmas associated with its abuse, and gaps in regulatory frameworks to mitigate these risks. In response to these issues, this study emphasizes the need to use AI-based solutions, including predictive analysis, anomalies identification, and collaborative threat intelligence, to proactively address these challenges. This comprehensive analysis aims to provide a deeper understanding of the role of AI in cyberattacks and promote innovative approaches to protect the integrity of digital ecosystems.

Keywords: Security Threats, Machine Learning (ML), Cyber Attacks – Artificial Intelligence (AI), Adversarial Artificial Intelligence.