Comparative Studies in Jurisprudence, Law, and Politics

Legal Analysis and Legislative Challenges of Artificial Intelligence in the Prevention and Counteraction of Computer Crimes and the Enhancement of Cybersecurity

1. Vahid Abdollahpour*: MA, Department of Criminal Law and Criminology, Damghan University, Damghan, Iran. Email: vahid.abdollahpour1367@gmail.com (Corresponding Author)

ABSTRACT

Computer crimes are among the prominent issues in the field of criminal law, and the importance of combating them—particularly in today's digital world—is becoming increasingly evident. In this regard, artificial intelligence (AI) is remarkably influential and can lead to profound transformations in computer technologies and cyberspace. The primary aim of this article is to examine how AI can play an effective role in countering computer crimes and improving cybersecurity performance. Based on research findings, given the limitations of human capabilities as well as the significant advancements in intelligent computer viruses and worms, the design of intelligent systems using advanced sensors and algorithms to combat computer crimes can be an effective step. One of the areas in which AI can strengthen cybersecurity is in the detection and identification of crimes. This technology, in addition to identifying threats, is also capable of preventing the occurrence of computer crimes. AI can enhance cybersecurity in various domains, including the detection of cyberattacks and threats such as botnets, the identification of spam and phishing attacks, combating fake accounts, protecting user information and data, authentication processes and fraud detection in authentication, monitoring user activities, and ensuring application security. Collectively, these measures can play a significant role in improving the safety of cyberspace. However, the use of AI to enhance cybersecurity also faces challenges. One of the major issues is the lack of transparency in AI decision-making processes, which may raise legal and ethical questions and challenges. Additionally, the requirement for large volumes of data to train algorithms and the necessity of human participation in decision-making processes are among the other challenges of this technology. Furthermore, the influence of various factors and complex variables in AI systems may cause difficulties in analyzing and predicting security threats. Ultimately, this article examines the legal dimensions and legislative challenges associated with AI in the prevention and counteraction of computer crimes and the enhancement of cybersecurity, and it proposes possible solutions to address these challenges.

Keywords: *Artificial Intelligence, Computer Crimes, Cybersecurity, Phishing, Network Security*

How to cite: Abdollahpour, V. (2025). Executive Guarantees for Compliance with International and National Regulations in Virtual Platforms Regarding the Commission of Crimes and the Dissemination of Criminal Content. *Comparative Studies in Jurisprudence, Law, and Politics*, *7*(4), 1-16.

Submit Date: 03 December 2024
Revise Date: 22 January 2025
Accept Date: 30 January 2025
Publish Date: 12 August 2025

پژوهش‌های تطبیقی فقه،
حقوق و سیاست

تحلیل حقوقی و چالش‌های قانونی هوش مصنوعی در پیشگیری و مقابله با جرائم رایانه‌ای و تقویت امنیت سایبری

۱. وحید عبداله پور*: کارشناس ارشد، گروه حقوق جزا و جرم شناسی، دانشگاه دامغان، دامغان، ایران. پست الکترونیک: vahid.abdollahpour1367@gmail.com (نویسنده مسئول)

چکیده

جرائم رایانه‌ای از جمله مسائل برجسته در حوزه حقوق کیفری به شمار می‌روند و اهمیت مقابله با آن‌ها به ویژه در دنیای دیجیتال امروز، روز به روز بیشتر نمایان می‌شود. در این راستا، هوش مصنوعی به‌طور چشمگیری تاثیرگذار است و می‌تواند به تحول و تغییرات شگرفی در فناوری‌های کامپیوتری و فضای مجازی منجر شود. هدف اصلی این مقاله بررسی این موضوع است که چگونه هوش مصنوعی می‌تواند در مقابله با جرائم رایانه‌ای و بهبود عملکرد امنیت سایبری نقش مؤثری ایفا کند. بر اساس نتایج تحقیقات، با توجه به محدودیت‌های موجود در توانایی‌های انسانی و همچنین پیشرفت‌های گسترده‌ای که در زمینه ویروس‌ها و کرم‌های رایانه‌ای هوشمند مشاهده می‌شود، طراحی سیستم‌های هوشمند با استفاده از سنسورها و الگوریتم‌های پیشرفته برای مقابله با جرائم رایانه‌ای می‌تواند گام مؤثری باشد. یکی از حوزه‌هایی که به کمک هوش مصنوعی می‌توان امنیت سایبری را تقویت کرد، شناسایی و کشف جرائم است. این فناوری علاوه بر شناسایی تهدیدات، قادر به پیشگیری از وقوع جرائم رایانه‌ای نیز می‌باشد. هوش مصنوعی در زمینه‌های مختلفی می‌تواند موجب بهبود امنیت سایبری شود. از جمله این موارد می‌توان به شناسایی حملات و تهدیدات سایبری مانند بات‌نت‌ها، شناسایی هرزنامه‌ها و حملات فیشینگ، مقابله با حساب‌های جعلی، حفاظت از اطلاعات و داده‌های کاربران، فرآیندهای احراز هویت و تشخیص تقلب در احراز هویت، نظارت بر فعالیت‌های کاربران و امنیت برنامه‌ها اشاره کرد. این اقدامات به‌طور کلی می‌توانند نقش بسزایی در افزایش ایمنی فضای سایبری ایفا کنند. با این حال، استفاده از هوش مصنوعی برای بهبود امنیت سایبری با چالش‌هایی نیز روبه‌رو است. یکی از مشکلات عمده، عدم شفافیت در فرآیندهای تصمیم‌گیری هوش مصنوعی است که ممکن است موجب بروز سوالات و چالش‌هایی در زمینه حقوقی و اخلاقی شود. همچنین، نیاز به داده‌های حجیم برای آموزش الگوریتم‌ها و همچنین ضرورت مشارکت انسانی در فرآیندهای تصمیم‌گیری از دیگر چالش‌های این فناوری محسوب می‌شوند. علاوه بر این، تاثیر عوامل مختلف و متغیرهای پیچیده در سیستم‌های هوش مصنوعی می‌تواند باعث بروز مشکلاتی در تحلیل و پیش‌بینی تهدیدات امنیتی گردد. در نهایت، این مقاله به بررسی ابعاد حقوقی و چالش‌های قانونی مرتبط با هوش مصنوعی در زمینه پیشگیری و مقابله با جرائم رایانه‌ای و تقویت امنیت سایبری پرداخته و به راهکارهای ممکن برای رفع این چالش‌ها اشاره خواهد کرد.

واژگان کلیدی: *هوش مصنوعی، جرائم رایانه‌ای، امنیت سایبری، فیشینگ، امنیت شبکه*

نحوه استناددهی: عبداله پور، وحید. (۱۴۰۴). ضمانت اجرایی برای رعایت ضوابط جهانی و ملی در پلتفرم‌های مجازی در قبال ارتکاب جرم و انتشار محتوای مجرمانه. *پژوهش‌های تطبیقی فقه، حقوق و سیاست*، *۷*(۴)، ۱۶-۱.

تاریخ ارسال: ۱۳ آذر ۱۴۰۳
تاریخ بازنگری: ۳ بهمن ۱۴۰۳
تاریخ پذیرش: ۱۱ بهمن ۱۴۰۳
تاریخ چاپ: ۲۱ مرداد ۱۴۰۴

مقدمه

تکامل چشمگیر فناوری‌های اطلاعات و ارتباطات، به ویژه اینترنت، به طور گسترده‌ای بر سازمان‌ها و زندگی اجتماعی تأثیر گذاشته است. این پیشرفت‌ها امکانات زیادی برای بهبود ارتباطات، تسهیل در تبادل اطلاعات و انجام فعالیت‌های مختلف به صورت آنلاین فراهم کرده است. با این حال، در کنار مزایای فراوانی که این تحولات به همراه داشته، چالش‌های جدی نیز در حوزه امنیت داده‌ها در فضای مجازی به وجود آمده است. یکی از این چالش‌ها، پیچیدگی‌های مربوط به پیگیری و مقابله با جرایم مجازی مانند سرقت‌های سایبری و کلاهبرداری‌های اینترنتی است که از ویژگی‌های فضای دیجیتال نشأت می‌گیرد. در دنیای دیجیتال امروز، فعالیت‌های اقتصادی، اجتماعی، و حتی فرهنگی به صورت عمده به فضای اینترنت منتقل شده است. این تغییرات باعث شده که میزان وابستگی به فناوری‌های اطلاعاتی و ارتباطی به طور چشمگیری افزایش یابد. اما این وابستگی موجب شده که تهدیدات امنیتی هم افزایش یابد. سرقت‌های اطلاعاتی، کلاهبرداری‌های آنلاین، حملات سایبری و دسترسی‌های غیرمجاز به داده‌ها، تهدیدهایی هستند که به طور مستمر سازمان‌ها، شرکت‌ها و حتی کاربران فردی را در معرض خطر قرار می‌دهند. یکی از مشکلات عمده در پیگیری این نوع جرایم، ماهیت غیرقابل لمس و نامرئی فضای مجازی است. برخلاف جرایم سنتی که اغلب در دنیای فیزیکی رخ می‌دهند و شواهد مادی قابل مشاهده‌ای از آن‌ها وجود دارد، جرایم سایبری به طور عمده در فضای دیجیتال اتفاق می‌افتند و این ویژگی، پیگیری آن‌ها را دشوارتر می‌کند. مثلاً در سرقت‌های اینترنتی، ممکن است هکرها از روش‌های پیچیده‌ای برای مخفی کردن ردپاهای دیجیتال خود استفاده کنند، به گونه‌ای که شناسایی و دستگیری آن‌ها برای نیروهای امنیتی و قضائی دشوار باشد. این مشکلات به ویژه زمانی بحرانی‌تر می‌شود که کلاهبرداران و مجرمان سایبری از ابزارهای پیچیده‌ای مانند فیشینگ، حملات DoS[1] یا بدافزارها برای فریب و سوءاستفاده از اطلاعات کاربران استفاده کنند. این حملات می‌توانند باعث سرقت اطلاعات حساس مانند شماره کارت‌های بانکی، کلمات عبور و حتی اطلاعات شخصی کاربران شوند. در نتیجه، پیگیری این جرایم نیازمند همکاری گسترده میان سازمان‌های دولتی و خصوصی، متخصصان امنیت سایبری و نیز ارتقاء آگاهی عمومی است. علاوه بر این، یکی دیگر از مشکلات اساسی در مقابله با جرایم سایبری، پیچیدگی‌های قانونی و حقوقی مرتبط با آن‌هاست (Calzavara et al., 2015). در بسیاری از موارد، مجرمان سایبری ممکن است از مرزهای جغرافیایی عبور کنند و فعالیت‌های خود را در کشورهای مختلف انجام دهند. این موضوع باعث می‌شود که پیگیری این جرایم به فرآیندی پیچیده و زمان‌بر تبدیل شود، زیرا نیاز به همکاری بین‌المللی و تعامل میان سیستم‌های قضائی مختلف وجود دارد. در مجموع، امنیت داده‌ها در فضای مجازی به یکی از مهم‌ترین چالش‌های عصر حاضر تبدیل شده است. هرچند که پیشرفت‌های زیادی در زمینه‌های فناوری امنیتی و روش‌های مقابله با جرایم اینترنتی انجام شده، اما هنوز هم این تهدیدات به طور کامل از میان برداشته نشده است. به همین دلیل، توجه به مسائل امنیت سایبری و اتخاذ تدابیر مناسب برای مقابله با این جرایم ضروری به نظر می‌رسد. در فضای مجازی، مرزهای جغرافیایی دیگر اهمیت چندانی ندارند و این امر باعث می‌شود مجرمان از هر نقطه‌ای از دنیا اقدام کنند. به همین دلیل، حملات سایبری به سازمان‌ها روز به روز پیچیده‌تر و چالش‌برانگیزتر می‌شود، به طوری که مقابله با آنها نیازمند استفاده از هوش مصنوعی یا تکنولوژی‌های پیشرفته است (Kleinmann & Wool, 2016).

[1] Denial of Service

.همانطور که جامعه ما بیشتر به هم متصل و از نظر فناوری پیشرفته‌تر می‌شود نقش راه حلهای امنیتی اهمیت بیشتری خواهد داشت. با این حال چالش ایمن سازی سیستم‌ها و جامعه ما که به این سیستم‌ها متکی است با تهدید مواجه است. از این رو طراحی راه حلهای امنیت سایبری کارآمدتر و مؤثرتر موضوعی است که همواره مورد علاقه است (Guan & Ge, 2017; Xiao, 2018).

از طرف دیگر، با توجه به حجم حملات سایبری، مداخله انسان در برابر این حملات برای دادن عکس‌العمل مناسب، کافی نیست، زیرا بیشتر این حملات، با استفاده از عامل‌های هوشمند مانند کرم‌ها و ویروس‌های کامپیوتری، صورت می‌گیرد. لذا مقابله با آن‌ها، نیاز به عامل‌های هوشمند دارد. این عامل‌های هوشمند وظیفه‌دارند: کل فرایند حمله را در زمان مناسب، مدیریت کنند، نوع حمله را شناسایی کنند، کشف کنند که هدف حمله چه بخشی از شبکه می‌باشد، مناسب‌ترین پاسخ به حمله، چه می‌باشد و چگونه جلوی حملات مشابه آن را بگیرند.علاوه بر این، مشکل جرائم رایانه‌ای، یک معضل جهانی شده است. در گذشته نه چندان دور، فقط افراد تحصیل‌کرده که با کامپیوتر و اینترنت سروکار داشتند در معرض این حملات بودند. اما امروزه با فراگیر شدن استفاده از کامپیوتر و اینترنت، هرکسی می‌تواند موردحمله سایبری قرار بگیرد. متأسفانه الگوریتم‌ها و برنامه‌های کامپیوتری که قبلا برای شناسایی این حملات، مورداستفاده قرار می‌گرفتند، دیگر کارا نیستند زیرا ویروس‌ها بسیار هوشمند شده‌اند و به‌محض ورود به یک سیستم، خود را با شرایط آن سیستم وفق می‌دهند و پیدا کردن آن‌ها ممکن نیست ما نیاز به استفاده از روش‌های یادگیری ماشین و هوش مصنوعی برای مقابله با این حملات داریم که هم انعطاف‌پذیر هستند و هم به‌سرعت با شرایط هر محیطی خود را سازگار می‌کنند. بر اساس آنچه گفته شد سوال مقاله بدین شکل قابل طرح است که هوش مصنوعی چه نقشی می‌تواند در مقابله با جرایم رایانه‌ای و بهبود عملکرد امنیت سایبری ایفا نماید؟ این مقاله توصیفی تحلیلی است و با استفاده از روش کتابخانه‌ای به بررسی سوال مورد اشاره پرداخته است. به منظور بررسی سوال مورد اشاره ابتدا نقش هوش مصنوعی در مقابله با جرایم رایانه‌ای بررسی شده و در ادامه از نقش آن در تحقق امنیت سایبری بحث می‌شود. در نهایت به محدودیت‌های هوش مصنوعی در بهبود امنیت سایبری پرداخته شده است.

۱. بکارگیری هوش مصنوعی برای مقابله با جرایم رایانه‌ای

توسعه سیستم‌های کامپیوتری، از یک‌طرف تأثیر مثبت قابل‌توجهی روی بهبود کیفیت زندگی بشر داشته از طرف دیگر مشکلات جدیدی مانند جرائم پیشرفته را با خود به ارمغان آورده است. جرائم معمولی مانند دزدی و کلاهبرداری، با بکارگیری فناوری اطلاعات، شکل جدیدی به خود گرفته‌اند که به آن "جرائم رایانه‌ای" می‌گویند که با پیشرفت فناوری، این جرائم نیز هرروز پیشرفته‌تر می‌شوند. از طرف دیگر.روند جهانی‌سازی و کمرنگ شدن نقش مرز بین کشورها، باعث شده که کشف، رصد، پیگیری و ممانعت از رخداد آن‌ها، هرروز سخت‌تر شود. (Gordon & Ford, 2006). اکثر جرائم سایبری که امروزه صورت می‌گیرند به‌سادگی نشان می‌دهند که جرائم، از فضای حقیقی به فضای مجازی، مهاجرت کرده‌اند به‌عبارت‌دیگر جرائم قدیمی با شیوه جدیدی رخ می‌دهند (Brenner, 2010). در این قسمت به بررسی بکارگیری هوش مصنوعی برای مقابله با جرایم رایانه‌ای پرداخته می‌شود.

۱-۱. نقش هوش مصنوعی در کشف جرائم ارایانه‌ای

هوش مصنوعی به‌عنوان یک علم جدید، برای اولین بار در سال ۱۹۵۶ ظاهر شد. هدف از آن ، ایجاد ماشین‌های هوشمند و حل مسائلی بود که بدون استفاده از درجاتی از هوشمندی، قابل‌حل نبودند. به‌عنوان‌مثال برای مقابله با جرائم رایانه‌ای، ما به سیستمی نیاز داریم که تا حدودی هوشمند باشد و بتواند از روی مقدار زیادی اطلاعات، تصمیمات درستی بگیرد. در مبحث هوش مصنوعی، با عامل سروکار داریم که تعریف

آن به‌صورت یک «موجود خودکار» که محیط اطراف خود را درک می‌کند. است. عامل دارای یک سیستم خود تصمیم‌گیرنده درونی است و عملی که انجام می‌دهد دارای تأثیر روی محیط اطراف است .هوش مصنوعی از تکنولوژی‌های متعدد برای حل مسائل استفاده می‌کند مانند: شبکه عصبی مصنوعی، منطق فازی محاسبات تکاملی، یادگیری ماشین، سیستم‌های ایمنی مصنوعی و کرم‌های مصنوعی. این فناوری‌ها توانایی تصمیم‌گیری انعطاف‌پذیر برای محیط‌های پویا مانند دایره امنیت و جرم را دارند. اکثر این روش‌ها، با الهام از طبیعت، از سیستم‌های بیولوژیکی به‌خوبی تقلید می‌کنند و مانند آن‌ها توانایی یادگیری، به خاطر سپردن، شناخت، طبقه‌بندی و پردازش اطلاعات رادارند.

در سال ۲۰۰۶ یک سیستم چندعاملی برای کشف کرم‌های کامپیوتری در محیط اینترنت شهری، طراحی گردید. در این محیط کرم‌های زیادی منتشر می‌شوند که پهنای باند زیادی از شبکه را به هدر می‌دهند و باعث ازکارافتادن مسیریاب‌ها می‌شوند (Gou et al., 2006). همچنین در سال ۲۰۰۶ سیستم دیگری بر اساس سیستم‌های چندعاملی طراحی شد که برای مقابله با حملات توزیعی DoS بسیار مناسب بود. این سیستم با استفاده از زبان پرولوگ پیاده‌سازی گردید و بدون دخالت انسان، عملکرد مناسبی داشت (Phillips et al., 2006). در سال ۲۰۰۷ چارچوب یک سیستم مقابله با حملات سایبری طراحی شد که در آن تعداد از عامل‌های هوشمند در تعامل با یکدیگر هستند و به‌مرورزمان، با توجه به شرایط شبکه و شدت و سختی حملات ، ساختار، پیکربندی و رفتار خود را تغییر می‌دهند (Kotenko & Ulanov, 2007).

۱-۲. نقش هوش مصنوعی در پیشگیری از جرائم رایانه‌ای

سیستم‌های ایمنی مصنوعی مانند سیستم ایمنی بیولوژیکی، برای نگه‌داشتن پایداری سیستم در یک محیط متغیر، طراحی‌شده‌اند. در سال ۲۰۰۷ یک سیستم مبتنی بر سیستم‌های ایمنی مصنوعی برای کشف هرزنامه‌ها، طراحی گردید (Sirisanyalak & Sornil, 2007). یکی از روش‌های انتشار ویروس‌ها و کرم‌های کامپیوتری، استفاده از هرزنامه و سوءاستفاده از اعتماد کاربران به ایمیل‌ها می‌باشد. در پروژه دیگری، با بررسی مدل‌های مختلف سیستم‌های ایمنی مصنوعی تئوری خطر برای اولین بار مطرح گردید. در این پروژه از نقشه‌های خودسازمانده برای پاسخ‌گویی به خطرات در شبکه بی‌سیم استفاده گردید (Lebbe et al., 2007). .

در سال ۲۰۱۲ یک سیستم کشف نفوذ مبتنی بر الگوریتم ژنتیک قانون‌گرا، طراحی شد. این سیستم برای بهبود امنیت، اطمینان، تجمیع و در دسترس بودن منابع شبکه، ارائه گردید. سیستم ارائه‌شده، از مجموعه‌ای از قوانین طبقه‌بندی که از داده‌های شبکه به‌دست‌آمده‌اند، تشکیل‌شده است و از درجه اطمینان شبکه به‌عنوان تابع پرارزش به‌منظور ارزیابی هر قانون، استفاده می‌کند (Ojugo et al., 2012). از قوانین فازی به‌منظور طبقه‌بندی انواع حملات استفاده می‌شود و الگوریتم ژنتیک کمک می‌کند تا قوانین فازی مناسب پیدا شوند.

۲. بکارگیری هوش مصنوعی برای بهبود عملکرد امنیت سایبری

کاربردهای هوش مصنوعی برای امنیت سایبری به طور کلی موفقیت آمیز بوده است. در مقابله با جرائم اینترنتی کمک‌های قابل توجهی انجام شده است. عمدتاً موضوعات مرتبط با سیستم‌های تشخیص و پیشگیری از نفوذ. سیستم‌های جدید بهبودهایی را نسبت به سیستم‌های قبلی نشان داده‌اند (Tsai et al., 2009).

فناوری‌های متداول پیشگیری امنیت سایبری از الگوریتم‌های ثابت و دستگاه‌های فیزیکی (مانند حسگرها و ردیاب‌ها) استفاده می‌کنند، بنابراین در مهار تهدیدهای جدید فضای مجازی بی اثر هستند. برای مثال، اولین نسل از سیستم‌های آنتی ویروس برای شناسایی ویروسها از طریق اسکن کردن بیت‌آن (Biton) طراحی شدند. فرض اساسی این مفهوم این است که یک ویروس همان ساختار و الگوی بیت را در همه نمونهها دارد. بنابراین این امضاها و الگوریتم‌ها ثابت هستند. اگرچه کاتالوگ امضاها به صورت روزانه به روز می‌شود (یا هر زمان که دستگاه به اینترنت

متصل باشد)، پیچیدگی و انتشار منظم بدافزارهای گسترده این روش را بی اثر می‌کند. با این حال، معرفی روش‌های بدون امضا که قادر به شناسایی و کاهش حملات بدافزار با استفاده از روش‌های جدیدتر مانند تشخیص رفتاری و هوش مصنوعی هستند، موثرتر است (Barth et al., 2012).

این نشان می‌دهد که پیشرفت در کاربردهای هوش مصنوعی امکان طراحی سیستم‌های نسبتاً کارآمد را فراهم کرده است که به طور خودکار فعالیت‌های مخرب را در فضای مجازی شناسایی و از آن جلوگیری می‌کند (Calzavara et al., 2015). آن‌ها برای حمایت از روش‌های فنآوری موجود به تصویب رسیده‌اند زیرا استانداردها و سازوکارهای موثری را برای کنترل و جلوگیری بهتر از حملات سایبری فراهم می‌کنند (Regev, 2009). علی رغم تمام مزایایی که هوش مصنوعی فراهم می‌کند، توسعه سریع آن باعث می‌شود تا محققان از بکارگیری کارآمدترین تکنیک و تأثیر آن بر امنیت فضای مجازی با مشکل مواجه شوند. هیچ ابهامی وجود ندارد که تصور عمومی در بین محققان امنیت اطلاعات و امنیت سایبری حاکی از آن است که هوش مصنوعی امنیت اطلاعات سازمانی را بهبود بخشیده است، اما با توجه به دانش ما، این ادعاها حدس و گمان است و از نظر تجربی اثبات نشده است. بیشتر مطالعات موجود نشان داده‌اند که چگونه نوآوری آنها از انتخاب روش‌های موجود بهتر عمل می‌کند یا نمونه‌ای از سیستم‌ها را بررسی کرده و عملکرد آنها را در مقایسه با آنها ارزیابی می‌کنند. در همه موارد، سطح تعصبات انتخاب نسبتاً زیاد است. بر این اساس، نیاز به یک ادبیات جمع بندی شده است که خلاصه‌ای از موضوعات، چالش‌ها و دستورالعمل‌های آینده تحقیق در این دامنه را ارائه دهد (Ali et al., 2017; Zhu & Tan, 2011). در این قسمت به بررسی بکارگیری هوش مصنوعی برای بهبود عملکرد امنیت سایبری پرداخته می‌شود.

۲-۱. امنیت داده و شبکه

سازمان‌ها برای به حداقل رساندن صدمات ناشی از حملات سایبری و جلوگیری از وقوع حوادث احتمالی و همچنین حفاظت از داده‌ها از فناوریهای نوینی مانند هوش مصنوعی استفاده می‌کنند. ابزارهای هوش مصنوعی می‌توانند داده‌ها را مورد بررسی قرار داده و داده‌های ناهنجار را که از دید انسان پنهان مانده است تشخیص دهند. یکی از کارکردهای هوش مصنوعی در زمینه امنیت داده شناسایی و جلوگیری از نشت داده می‌باشد نشت داده ممکن است از داخل و یا خارج از سازمان اتفاق افتد که در بسیاری از مواقع آسیب جدی به اطلاعات و دانش سازمانها وارد می‌نماید این امر می‌تواند به دلیل اتصال کاربران غیرمجاز به سیستم، سطح دسترسیهای نادرست به اطلاعات و یا معماری اشتباه شبکه باشد سیستم‌های مجهز به هوش مصنوعی با شناسایی نظارت و مراقبت از اطلاعات موجود در حافظه و همچنین مراقبت از دادههایی که در شبکه‌ها در حال جابه جایی هستند، امنیت دادهها را افزایش می‌دهد و در صورت وجود رفتار غیرعادی به طور خودکار هشدار داده و از ادامه آن رفتار جلوگیری می‌کند (Kowert, 2017).

با گسترده شدن شبکه‌های کامپیوتری ایجاد امنیت در شبکه از اهمیت به سزایی برخوردار است. متأسفانه افراد سودجو با دسترسی به اطلاعات مهم مراکز خاص و یا اطلاعات افراد دیگر به قصد اعمال نفوذ فشار و یا ایجاد بینظمی در سیستم‌ها، حمله به شبکه و سیستم‌ها را در پیش می‌گیرند. دو مدل یادگیری با ناظر و یادگیری بدون ناظر در حوزه هوش مصنوعی توانستند تأثیر بسزایی در تأمین امنیت شبکه بگذارند مدلهای یادگیری ماشین با ناظر بر اساس داده‌های برچسب دار، آموزش می‌بینند که مسائل طبقه بندی نیز در این گروه قرار می‌گیرند در این مسائل داده‌های ورودی بر اساس الگوی نهفته در آنها در یک دسته مشخص جای می‌گیرند؛ اما در یادگیری بدون ناظر به توسعه سیستم‌هایی پرداخته می‌شود که تنها با یافتن شباهت بین داده‌های بدون برچسب آنها را گروه بندی می‌نماید با توجه به این که داده‌ها برای این مدلها همگی یکسان

به نظر می‌رسند (بدون برچسب، هدف این نوع از الگوریتم‌ها درک رابطه بین دادهها و در نهایت گروه بندی آنها ست خوشه بندی، یکی از نمونه‌های رایج این نوع از مدلهاست. در ادامه کاربردهای هوش مصنوعی در امنیت شبکه شرح داده می‌شود (Bowman & Huang, 2019).

۲-۱-۱. تشخیص تهاجم

سیستم تشخیص تهاجم و) یا (نفوذ یک ابزار مؤثر جهت شناسایی و تشخیص هر گونه استفاده غیر مجاز سوءاستفاده و یا آسیب رسانی در شبکه را بر عهده دارد. «برای ایجاد امنیت کامل در یک سیستم کامپیوتری، علاوه بر دیوارهای آتش و دیگر تجهیزات جلوگیری از نفوذ سیستم‌های دیگری نیز نیاز است تا بتوان در صورت عبور نفوذگر از دیواره آتش آنتی ویروس و سایر تجهیزات امنیتی آنها را تشخیص داده و چاره‌ای برای مقابله با آن تعبیه نمود. به دلیل ماهیت غیر الگوریتمی روش‌های نفوذ در شبکه‌های کامپیوتری راهکارهای مطرح شده برای مقابله با ناهنجاریها نیز باید دارای ماهیت غیر الگوریتمی باشد. هوش مصنوعی به ویژه شبکه عصبی که جزء سیستم‌های یادگیرنده محسوب می‌شود، توانسته در زمینه شناسایی و جلوگیری از این ناهنجاریها تأثیر بسزایی بگذارد. به طور سنتی، فعالیت شناسایی نفوذ از طریق معرفی سیستم‌هایی معروف به سیستم‌های تشخیص نفوذ مدیریت می. شود. این سیستم‌ها به دو دسته سیستم تشخیص نفوذ تحت شبکه و سیستم‌های تشخیص نفوذ میزبان تقسیم می‌شوند که بر اساس عواملی از قبیل نوع و تعداد فرآیندهای در حال اجرا، رفتار کاربران ماژولهای سیستم عامل در هنگام راه اندازی سیستم و یا الگوی مصرف ترافیک به نظارت و یا کنترل شبکه‌های داخلی و یا خارجی می‌پردازند. با معرفی تکنیکهای هوش مصنوعی در زمینه امنیت سایبری، نوع سومی تحت عنوان سیستم تشخیص نفوذ مبتنی بر ناهنجاری مطرح گردید این دسته با استفاده از الگوریتمهای یادگیری با ناظر و یا بدون ناظر، یادگیری تقویتی و حتی الگوریتم‌های خوشه بندی حوزه داده کاوی رویدادهایی را که خارج از رفتار عادی سیستم اتفاق میافتد به عنوان ناهنجاری شناسایی می‌نماید» (Ghorbanpour Vishkasouqi & Mirazimi Tabalvandani, 2023).

۲-۱-۲. شناسایی بات نت‌ها

یکی از رایج‌ترین مشکلات در تشخیص ناهنجاری، شبکه مربوط به باتنتهاست. با توجه به خطر چنین شبکههای مخفی شناسایی بات نت‌ها جهت جلوگیری از فرسودگی شبکه و مصرف بیهوده منابع محاسباتی و برای جلوگیری از انتشار اطلاعات حساس نشت دادهها به خارج از شرکت از جمله فعالیتهای ضروری جهت برقراری امنیت داده‌ها و اطلاعات در سطح شبکه است بات نت‌ها از مجموعه‌ای از کامپیوترها و یا سیستم‌های آلوده تشکیل شده‌اند که توسط یک بدافزار هدایت می‌شوند. افرادی که بدافزارها را ایجاد کرده‌اند می‌توانند به جای این که به یک کامپیوتر آلوده به صورت ستقیم متصل شوند، از بات نت‌ها برای مدیریت خودکار حجم زیادی از این کامپیوترهای آلوده استفاده کنند که از طریق یک کانال فرمان و کنترل C&C به یکدیگر وصل شده‌اند یکی از مشکلات کشف بات نت‌ها شباهت زیاد ترافیک بات نت با ترافیک واقعی شبکه است بنابراین به راحتی و با استفاده از روش‌های سنتی نمی‌توان آنها را شناسایی نمود. الگوریتم‌های یادگیری ماشین در دو نوع با ناظر و بدون ناظر یکی از ابزارهای هوش مصنوعی در شناسایی بات نت‌ها هستند. شناسایی بات نت‌ها نمونه‌ای از مسائل طبقه بندی است که با هدف تعیین کلاس یک بسته یا توالی بسته‌ها به ترافیک بات نت‌ها و یا ترافیک معمولی مورد بررسی قرار می‌گیرند. از طرف دیگر جهت گروه بندی ترافیک با توجه به ویژگی‌های مشابه و شناسایی ترافیک‌های مشکوک می‌توان از مدلهای یادگیری ماشین بدون ناظر استفاده کرد (Bakhshi & Veisi, 2019).

۲-۲. امنیت ایمیل

تهدیدات امنیتی از ایمیل نیز به عنوان ابزاری برای حمله استفاده می. کند از آنجا که میزان ترافیک منتقل شده از این طریق به سیار زیاد است استفاده از روش‌های تشخیص خودکار که از الگوریتمهای یادگیری ماشین بهره می‌برند، امری ضروری در شناسایی هرزنامه و حملات فیشینگ است کاربردهای هوش مصنوعی در امنیت ایمیل در ذیل شرح داده شده است (Bakhshi & Veisi, 2019). امنیت ایمیل‌ها از دو طریق شناسایی هرزنامه و شناسایی حملات فیشینگ انجام می‌شود. «هرزنامه یا اسپم به پیام الکترونیکی اطلاق می‌شود که بدون درخواست گیرنده و برای افراد بی‌شمار فرستاده می‌شود. هرزنامه به نوعی سوء استفاده از سامانه انتقال پیام است. جهت شناسایی هرزنامه، استراتژی‌های هوش مصنوعی متفاوتی وجود دارد که یکی از رایج‌ترین و ساده‌ترین آن‌ها شبکه‌های عصبی است. از قابلیت‌های دیگری همچون ماشین‌های بردار پشتیبان (به‌ویژه برای اسپم‌های تصویری)، شبکه‌های بیز و همچنین استفاده از فناوری‌های پردازش زبان طبیعی می‌توان در این زمینه استفاده نمود» (Keshavarz & Hosseini, 2023).

۲-۳. حفاظت از حساب‌ها و اطلاعات کاربران و مقابله با حسابهای جعلی

«حفاظت از حساب‌های کاربری برای هر سازمانی به خصوص در شرایطی که دارای مسئولیت‌های قانونی در قبال اشخاص ثالث است، یک مأموریت حساس به شمار می‌رود. این امر با گسترش حساب‌های جعلی جهت سرقت اطلاعات محرمانه و حساس و یا ایجاد تشنج و بیان مطالب محرک و توهین‌آمیز در فضای سایبری بحرانی‌تر شده است. یکی از نقاط ضعف در محافظت از حساب کاربران، محافظت ضعیف از رمز عبور آن‌هاست. هر چند امروزه در راستای افزایش امنیت حساب‌های کاربران، راهکارهایی از جمله ارسال پیام و یا ایمیل به کاربر در حین اتصال سیستم دیگری به حساب مربوطه مطرح شده است. با این حال، این فعالیت‌ها، واکنشی هستند که با شناسایی دسترسی‌های غیرمجاز، سیستم در قالب هشدار، واکنشی را نشان داده و باعث بسته و یا معلق شدن حساب کاربر می‌گردد. این سیستم‌های هشدار واکنشی، معمولاً با مجموعه‌ای از محرک‌های پیش‌فرض و مرتبط با رویدادها فعال می‌شوند که برای تمامی کاربران یکسان هستند. به عبارت دیگر، این سیستم‌ها در شناسایی رفتار کاربران تلاشی نکرده تا بتوانند بر اساس الگوی رفتاری هر فرد عمل نمایند. علاوه بر این، سیستم‌های واکنشی، آینده را مشابه با گذشته در نظر می‌گیرند و توانایی سازگاری سریع با تغییرات را ندارند» (Ghorbanpour Vishkasouqi & Mirazimi Tabalvandani, 2023). اتخاذ یک رویکرد پیشگیرانه برای حفاظت حساب و اطلاعات کاربران می‌تواند مثمر ثمر واقع شود. در این جاست که هوش مصنوعی با استفاده از روش‌های مختلف داده کاوی و یادگیری ماشین جهت بهره برداری از داده‌های ساختاریافته و یا غیر ساختاری استخراج شده از منابع ناهمگون سازمان به کار می‌آید هوش مصنوعی با شروع تجزیه و تحلیل داده‌های گذشته قادر به نشان دادن الگوهای نهفته برآورد رفتارهای آینده کاربران و شناسایی به موقع تلاش‌های احتمالی برای کلاهبرداری است (Adekunle, 2019). مقابله با حساب کاربری جعلی از جمله فعالیتهای مستلزم نظارت است؛ زیرا گسترش حساب‌های جعلی باعث فراهم نمودن بستری نا امن برای انجام فعالیتهای نامنا سب از جمله فریب کاربران قانونی و استفاده از حساب کاربری آن‌ها می‌گردد نظارت بر حساب کاربری شبکه‌های اجتماعی مانند فیسبوک و توییتر که روزانه تعداد زیادی حساب کاربری ایجاد می‌شود کاری دشوار و زمانبر است که با گسترش کاربردهای هوش مصنوعی در امنیت، این فعالیت با دقت بالا و زمان کم قابل انجام است به عنوان، نمونه فیسبوک با به کارگیری سیستم جدیدی مبتنی بر یادگیری ماشین موسوم به "Deep Entity Classification" به شناسایی حسابهای جعلی می‌پردازد و از فعالیت آنها جلوگیری می‌کند.

این الگوریتم ۲۰۰۰۰ ویژگی متعلق به هر حساب و افراد وابسته به آن حساب را در نظر می‌گیرد و اصل بودن آن را بررسی می‌نماید. انجام این عمل به صورت دستی توسط انسان کاری تقریباً غیرممکن است (Adekunle, 2019).

۲-۴. احراز هویت و تشخیص فریب احراز هویت

احراز هویت فرآیندی است که طی آن درستی هویت یک فرد شناسایی و تأیید می‌گردد. بنابراین زمانی که کاربر بخواهد وارد سیستم شده و یا به منبعی دسترسی پیدا کند، ابتدا باید خود را اثبات نماید (Chang et al., 2016). احراز هویت به صورت روش‌های متفاوتی از قبیل سؤال‌های، امنیتی رمزهای عبور توکن ها، دستگاه‌های فیزیکی و ویژگی‌های بیومتریک انجام می‌گیرد در حال حاضر احراز هویت از طریق ویژگی‌های بیومتریک بسیار مطرح شده است منظور از ویژگی‌های بیومتریک در احراز هویت، ویژگی‌های فیزیکی منحصر به فرد مانند عنبیه چشم چهره اثر انگشت و صداست که از طریق آن می‌توان افراد را شناسایی و ردیابی نمود هوش مصنوعی با استفاده از تکنولوژیهای بینایی ماشین و تشخیص گفتار، تحول شگرفی در این زمینه ایجاد کرده است که می‌تواند با تلفیق با علم داده کاوی الگوهای رفتاری مانند نحوه تایپ مطالب و دست خط و یا الگوی امضا کردن را نیز شناسایی نماید و از آن برای صحه گذاری هویت افراد بهره ببرد. علاوه بر این از قابلیتهای دیگر سیستم‌های هوشمند می‌توان به موارد زیر اشاره کرد (Adekunle, 2019).

دسترسی کاربر به عنوان اولین خط دفاعی امنیت سایبری، این سیستم باید مدیریت احراز هویت دسترسی کاربر را تقویت، کند انواع رفتارهای استتاری را با دقت شناسایی کند، و تشخیص اشیاء غیرقانونی یا مخرب را شناسایی کند. قبل از عملیات سیستم باید از احراز هویت کاربران اطمینان حاصل کند در عین حال داده‌های کاربر باید محرمانه باشد تا از سایر رویدادهای خطرناک مانند جمع آوری مخرب اطلاعات کاربر جلوگیری شود. در فرآیند احراز هویت فعلی بر افزودن ویژگی‌های دیگری برای افزایش منحصر به فرد بودن فرآیند تطبیق رمز عبور تمرکز می‌کند تا احتمال عبور دیگران به عنوان کاربران قانونی را به حداقل برساند.

حالت نحوه تطبیق رمزهای عبور و افزودن سایر مشخصات کاربر برای اطمینان از امنیت احراز هویت دوگانه چالشی است که باید در احراز هویت حالت حل. شود به عنوان مثال، دستگاه‌های خودپرداز فعلی فقط از کدهای پین برای تأیید هویت استفاده می‌کنند که به تنهایی امنیت احراز هویت را تضمین نمی‌کند (Adekunle, 2019). با توجه به کاستیهای احراز هویت تک مرحله‌ای فناوری احراز هویت چندگانه مانند شوفان در نظر گرفته شده است برای رسیدن به این هدف از جنگل تصادفی استفاده کرد کورکماز نه تنها تطبیق رمز عبور را در سیستم احراز هویت رمز عبور انجام داد، بلکه صفحه کلید کاربر را با استفاده از برخی سبکها از طریق شبکه عصبی آموزش داد این سبکها شامل سرعت تایپ کاربر و سبک تایپ، ترکیب کلیدها و سایر جنبه‌ها بود. وانگ و فانگ یک تابع هسته با عملکردهای جهانی و محلی طراحی کردند و یک مکانیسم احراز هویت امنیتی شبکه ارتباطی تلفن همراه را بر اساس رگرسیون بردار پشتیبانی (SVR) ساختند، اما داده‌های کمتری در شبیه سازی استفاده شد. چانگ و همکاران از ماشین بردار پشتیبانی یک کلاس SVM یک کلاس) برای تشخیص الگوی پویایی ضربه زدن به کلید استفاده کرد و این الگو به دلیل هوش مصنوعی توجه گسترده‌ای را به خود جلب کرد (Ali et al., 2017). لو و همکاران از شبکه عصبی کانولوشن(CNN) یادگیری تقویتی و یادگیری انتقال برای ساخت یک طرح احراز هویت فیزیکی استفاده کردند. هدف آن محاسبات لبه موبایل بود و برای مقاومت در برابر حملات لبه‌های سرکش استفاده شد.

با پیشرفت تکنولوژی علاوه بر ایجاد راهکارهایی جهت تأمین امنیت بیشتر در فضای دیجیتال می‌توان ادعا نمود که به همان میزان حملات و تقلبهای سایبری نیز پیچیده‌تر و پیشرفته‌تر شده است. جعل در تصاویر در حوزهٔ Anti Spoofing شامل موارد متفاوتی از جمله جعل در

صحبت کردن و عدم رعایت الگوی موردنظر در صحبت می‌باشد. با ظهور فناوری دیپ فیک و مدلهایی مانند مدلهای مولد تخاصمی می‌توان تصاویری از افراد ایجاد کرد که هرگز حضور فیزیکی نداشته‌اند. و سیستم‌های مربوط به احراز هویت را فریب داد در این راستا، فرآیندی تحت عنوان صحت سنجی یا تشخیص زنده بودن مطرح شده که عبارت است از مجموعه‌ای از عملیات که تصاویر ویدئویی را مورد پردازش قرار داده و تشخیص می‌دهد که ویدئوی دریافت شده از فرد جعلی نبوده و مورد دستکاری عامدانه قرار نگرفته است. این مورد یکی از جالب‌ترین کاربردهای هوش مصنوعی در امنیت سایبری است که به نوعی این فناوری را در مقابل خود قرار می‌دهد (Adekunle, 2019).

۲-۵. امنیت برنامه

تأمین امنیت برنامه یکی دیگر از کاربردهای هوش مصنوعی در امنیت سایبری شامل ملاحظات امنیتی برای مواردی از قبیل برنامه‌های وب برنامه‌های کاربردی کلاینت و موبایل است که در طول، توسعه طراحی برنامه و پس از استقرار آن اتفاق میافتد. به طور کلی برنامه‌ها در همه مکانها قابلیت اجرا شدن دارند و دائماً در حال تغییر هستند به همین دلیل تأمین امنیت آنها د شوار است. هدف از امنیت برنامه جلوگیری از سرقت یا ربودن دادهها یا کدهای موجود در برنامه است (Singh, 2017).

نقض امنیت سایبری در لایه برنامه می‌تواند به دلایل مختلف از قبیل کدنویسی، ضعیف، انجام کم تست و ر شد سریع و چشمگیر تکنولوژیهای جدید مانند سرورهای ابری باشد که این لایه را در مقابل حمله مهاجمان آسیب پذیر کرده است. با این حال، حوزه‌های مختلف هوش مصنوعی از جمله سیستم‌های خبره و یادگیری ماشین توانسته‌اند در بهبود تجزیه و تحلیل و پیش بینی حملات امنیتی شناسایی نقاط آسیب پذیر برنامه و ارائه گزینه‌های اصلاحی جهت رفع عیوب کدگذاری تأثیرات قابل توجهی بگذارند به عنوان، نمونه مایکرو سافت با تکیه بر هوش مصنوعی ابزاری ارائه کرده که به توسعه دهندگان کمک می‌کند ایرادهای کدهایشان را با دقت ۹۹ درصد شناسایی کنند (Bakhshi & Veisi, 2019) این کار باعث می‌شود بخش بزرگی از امنیت برنامه در همان ابتدا و در مرحله طراحی و توسعه، تأمین گردد.

۳. محدودیت‌های هوش مصنوعی در بهبود امنیت سایبری

مانند هر چیز دیگری، استفاده از هوش مصنوعی در زمینه امنیت سایبری دارای معایبی بوده و با محدودیت‌هایی همراه است. به منظور ایجاد و حفظ یک سیستم هوش مصنوعی، سازمان‌ها به منابع و سرمایه گذاری‌های بیشتری نیاز دارند. علاوه بر این از آنجایی که سیستم‌های هوش مصنوعی با استفاده از مجموع داده‌ها آموزش می‌بینند، باید مجموعههای متمایز زیادی از کدهای بدافزار، کدهای غیرمخرب و ناهنجاریها را بهدست آورد. دستیابی به این همه مجموعه داده‌ها زمان بر است و نیاز به سرمایه گذاریهایی دارد که اکثر سازمانها قادر به پرداخت آن نیستند . بدون حجم عظیمی از داده‌ها و رویدادها، سیستم‌های هوشمصنوعی می‌توانند نتایج نادرست و یا مثبت کاذب ارائه دهند و دریافت داده‌های نادرست از منابع غیرقابل اعتماد حتی می‌تواند نتیجه معکوس داشته باشد. نکته منفی دیگر این است که مجرمان سایبری می‌توانند از هوش مصنوعی برای تجزیه و تحلیل بدافزار خود و انجام حملات پیشرفتهتر استفاده کنند که ما را به نقطه بعدی می‌رساند. مشکل تضمین امنیت اطلاعات محرمانه یکی از کلیدهای همه موضوعات اقتصاد دیجیتال از جمله مشکل امنیت سایبری با استفاده از هوش مصنوعی است. جامعه جهانی نگران استفاده از هوش مصنوعی برای مقاصد جنایی است (Bahramizadeh & Sadeghzadeh, 2022). از طرفی آیا هوش مصنوعی می‌تواند تمام رویدادهای نامطمئن را شناسایی کند؟ پاسخ قطعاً منفی است. این فناوری جدید به عنوان یک «شمشیر دولبه علاوه بر مزیتهای فراوان دارای کاستیهای خاص خود نیز می‌باشد این بخش عواملی را مورد بحث قرار می‌دهد که مدل هوش مصنوعی را در زمینه امنیت سایبری ناصادق می‌کند.

۳-۱- تداخل داده‌های گیج کننده

چه مقدار تداخل می‌تواند هوش مصنوعی را فریب دهد؟ شاید یک پیک سل کافی باشد. آزمایش سو و همکاران در سال ۲۰۱۹ نشان داد که تنها تغییر یک پیکسل در تصویر می‌تواند منجر به طبقه بندی اشتباه شبکه عصبی شود. همانطور که از این مثال مشاهده می‌شود، زمانی که داده‌ها «آلوده شوند»، فرصتی برای تقلب در سیستم هوش مصنوعی وجود دارد که منجر به وضعیت ناامن شبکه می‌شود.

۳-۲. عدم شفافیت در فرآیند تصمیم گیری هوش مصنوعی

در فرآیند تصمیم گیری هوش مصنوعی همه شرکت کنندگان از جمله برنامه نویسان نمی‌دانند که چرا و چگونه مدل هوش مصنوعی نتایج نهایی تصمیم گیری را ارائه می‌دهد، یعنی فرآیند تصمیم گیری هوش مصنوعی فاقد شفافیت است. مدل هوش مصنوعی شبیه جعبه سیاه است در فرآیند ایجاد و خودسازی می‌تواند پیکربندی و تنظیم خودکار پارامترها را بدون دخالت بیش از حد کارکنان محقق کند و در نتیجه موجب صرفه جویی در منابع انسانی شود. با این وجود، در عین حال مشکل این است که فرآیند تصمیم گیری آن دشوار است که به وضوح توضیح داده شود. اگرچه مدل هوش مصنوعی می‌تواند به دقت بالایی دست یابد، اما همه آزمایشها در مجموعه آزمایشی پیاده سازی می. شوند بنابراین هنگام مواجهه با رویدادهای ناشناخته، اینکه آیا مدل هوش مصنوعی می‌تواند به چنین دقت بالایی دست یابد باید تأیید شود. هنگامی که به نتایج تصمیم گیری ارائه شده توسط مدل هوش مصنوعی اعتراضاتی وجود دارد تو ضیح فرآیند تصمیم گیری دشوار است بنابراین برخی از افراد نسبت به نتایج تصمیم گیری بدبین خواهند بود این برای قضاوت سریع وضعیت شبکه مفید نخواهد بود و یا حتی باعث عواقب جبران ناپذیری خواهد شد. برخی از تیمهای تحقیقاتی شروع به انجام تحقیقات عمیق در مورد این موضوع کرده‌اند (Schlegel, 2019).

۳-۳. نیاز به داده‌های بالا

در حال حاضر، مدل‌های هوش مصنوعی اساساً به داده‌های زیادی برای تکمیل آموزش نیاز دارند. قبل از استفاده از داده‌ها، آن‌ها ممکن است عملیاتی را انجام دهند که عمدتاً شامل یک سری مراحل مانند کاهش نویز، داده‌ها، نرمال سازی، پر کردن مقادیر گم شده و از دست رفته و غیره است. اگر از روش نظارتی استفاده می‌شود لازم است دادهها را به صورت دستی برچسب گذاری کرد با این حال به دلیل ناهمگونی شدید فضای، سایبری ساختارهای سایبری مختلف ممکن است رویدادهای پرخطر متفاوتی را ایجاد کنند و این رویدادها دارای ویژگی‌های ناگهانی هستند بنابراین هر رویداد احتمالی پرخطر را نمی‌توان قبل از طراحی مدلها تخمین زد و همچنین نمی‌توان این رویدادهای پرخطر را از قبل تحلیل و برچسب گذاری کرد. در همین حال مدلهای هوش مصنوعی تقاضای بالایی برای دادهها دارند که ممکن است نتوانند به موقع بررسی کنند.

۳-۴. ضرورت مشارکت انسان با هوش مصنوعی

طراحی و اجرای هوش مصنوعی به ویژه شبکه‌های عصبی سعی در تقلید از مغز انسان دارد. هدف آن دستیابی به همان شیوه تفکر انسان با استفاده از نورونهای متصل است. متأسفانه هوش مصنوعی برای دستیابی به یادگیری به تعداد زیادی نمونه نیاز دارد. بدون توانایی استدلال، مدل نهایی بعد از آموزش برای ما پیچیده است تا بفهمیم چگونه تصمیم می‌گیرد. اگرچه برخی تلاشها برای یافتن توضیح هوش مصنوعی انجام شد اما این کار هنوز در مرحله اولیه است بنابراین، برای دستیابی به استفاده کارآمد از هوش تکیه بر این ابزارها بدون مشارکت انسانی کافی نیست (Nunes et al., 2015).

هوش مصنوعی نقش مهمی در پیشگیری و تشخیص رفتارهای پرخطر شبکه ایفا می‌کند اما عوامل متعددی وجود دارد که در قضاوت صحیح آن اختلال ایجاد کنند هوش مصنوعی برای کمک به متخصصان امنیتی در این زمینه است، نه جایگزین کردن آن‌ها بنابراین همچنان لازم است متخصصان مربوطه مداخله کنند و از دانش شبکه مربوطه برای قضاوت حرفه‌ای در مورد فرم فعلی شبکه استفاده کنند.

در حال حاضر نوع جدیدی از هوش مصنوعی در حال توسعه است که انسان در حلقه نامیده می‌شود. در سال ۲۰۱۷، آژانس پروژه‌های تحقیقاتی پیشرفته دفاعی دارپا چالش روباتیک دارپا را طراحی کرد. در ژانویه ۲۰۱۹، دارپا پروژه هوش مصنوعی به نام KAIROS را برای پیاده سازی سیستمی منتشر کرد که می‌تواند رویدادها را شناسایی کرده و توجه انسانها را به خود جلب کند و همچنین در ماه مه، راه اندازی پروژه ACE را برای توسعه قابلیت نبرد هوایی مشترک بین انسان و ماشین اعلام کرد در ماه نوامبر وزارت دفاع ایالات متحده گزارشی در مورد جنگندههای مکانیکی و ادغام انسان و ماشین دریافت کرد که در آن مبارزان سایبورگ برای جنگهای آینده ساخته می‌شدند در زمینه نظامی نیز تحقیقات این فناوری جدید آغاز شده است که اهمیت آن را نیز منعکس می‌کند. تکنیک انسان در حلقه می‌تواند خرد انسانی و هوش ماشینی را ترکیب، کند که روشی مهم برای درک مزایای مکمل انسان و ماشین است. هوش مصنوعی می‌تواند تعداد زیادی از داده‌ها را به سرعت پردازش کند و جلوه تشخیص خوبی برای صحنه‌های خاص دارد. اما ممکن است مختل شود و و ضعیت جدید را به درستی قضاوت نکند در مقایسه با ماشینها، از سانها انعطاف پذیرتر هستند و می‌توانند در مواجهه با تغییرات جدید در شبکه سریعتر قضاوت، کنند اما همچنین به ماشینهایی برای ارائه نقش کمکی نیاز دارند یادگیری ماشین، تعاملی که در هوش مصنوعی مورد استفاده قرار گرفت در امنیت سایبری نیز تجسم یافته است(Santhanam GR, et al,2017:209-230). افزودن این ایده به آگاهی، موقعیتی ضمن دستیابی به تجسم قابلیت اطمینان سیستم را نیز افزایش می‌دهد. بنابراین استفاده از هوش مصنوعی و انسان در حلقه در امنیت سایبری قابلیت‌های مدل‌ها را بیشتر خواهد کرد.

نتیجه‌گیری

استفاده از هوش مصنوعی در حوزه امنیت سایبری امیدوار کننده است و سیستم‌های تشخیص و پیشگیری از نفوذ در حال پیشرفت است. هوش مصنوعی باعث کاهش پیچیدگی محاسباتی و کاهش زمان آموزش مدل شده است. همچنین مشاهده شد که یک انحراف قابل توجهی در دامنه وجود دارد. علاوه بر این، محققان روی الگوریتم‌های کمتری تمرکز کرده‌اند و به همین ترتیب الگوریتم‌های جدید محبوب نیستند. این امر به عنوان یک چالش و همچنین فرصتی برای محققان به وجود می‌آید. اعتقاد بر این است که برنامه‌های هوش مصنوعی همچنان فرصت‌هایی برای امنیت سایبری ارائه می‌دهند .یکی از کارهایی که در این حوزه باید انجام گیرد پیدا کردن تئوری "توابع تسهیلاتی گروهی" برای اجازه دادن به عامل‌ها برای تصمیم‌گیری گروهی است. استفاده از الگوریتم‌های یادگیری بدون نظارت برای ساخت سیستم‌های ترکیبی "کشف نفوذ غیرمعمول". ترکیب کلیه تکنیک‌های هوش مصنوعی برای ساخت ویروس‌یاب‌های جدید، پیشنهادات دیگر در این زمینه می‌باشند. با توجه به محدودیت‌های انسانی و این حقیقت که ویروس‌ها و کرم‌های رایانه‌ای هوشمند و انعطاف‌پذیر شده‌اند، طراحی عامل‌های با سنسور هوشمند بسیار ضروری می‌باشد. این عامل‌ها توانایی کشف، ارزیابی و پاسخ‌دهی به حملات سایبری در زمان مناسب را باید در خود داشته باشند. به‌کارگیری تکنیک‌های هوش مصنوعی در حوزه دفاع سایبری نیاز به برنامه‌ریزی و مطالعه و تحقیق دارد. پیش بینی در زمینه‌های مربوط به جرم از طریق مطالعه نظام‌مند روندها، آثار و عوارض آن‌ها را درآینده بررسی و تبیین کرده و براساس آنها تمهیداتی را برای جلوگیری از رخدادها یا کاهش تبعات در نظر می‌گیرد. حال با توجه به این که یکی از کاربردی ترین وتخصصی ترین استفاده‌های هوش مصنوعی،

پیش‌بینی و تصمیم‌گیری در مقابل داده‌ها و قوانینی است که ازقبل برای آن تعریف شده است، این علم می‌تواند در زمینه پیشگیری، به خصوص پیشگیری انتظامی از جرم کمک شایانی به نهادهایی کند که وظیفه پیشگیری انتظامی ازجرم را برعهده دارند، ولی مشروط براین که داده‌هایی که به این عامل وارد می‌شوند، و همچنین قوانینی که برای آن تعریف می‌شود، به صورت واقعی، دقیق و منطقی باشد تا بتوان شاهدذ خروجی کارآمد و مناسبی بود. هوش مصنوعی با استفاده از شبکه‌های عصبی مصنوعی، طراحی عامل یادگیر، منطق فازی، پردازش سیگنال نقش به سزایی را در پیشگیری انتظامی ازجرم ایفا می‌کند و ما می‌توال با استفاده مناسب ازخروجی این عامل‌ها و همچنین با دسته بندی و در اختیار قرار دادن این اطلاعات به نهادهای ذیربط گامی نو و اساسی در راستای پیشگیری انتظامی ازجرایم برداشت.

تعارض منافع

در انجام مطالعه حاضر، هیچ‌گونه تضاد منافعی وجود ندارد.

مشارکت نویسندگان

در نگارش این مقاله تمامی نویسندگان نقش یکسانی ایفا کردند.

حامی مالی

این پژوهش حامی مالی نداشته است.

EXTENDED SUMMARY

The rapid advancement of information and communication technologies, especially the Internet, has brought significant transformations to social, economic, and cultural spheres, but it has also introduced profound security challenges in cyberspace. Computer crimes—ranging from identity theft and online fraud to sophisticated cyberattacks—have escalated in scale and complexity, posing significant risks to individuals, organizations, and governments. Unlike conventional crimes, cybercrimes operate in a borderless, intangible digital space, making detection, prosecution, and deterrence considerably more difficult. For example, the anonymity and global reach of cyberspace enable perpetrators to operate across jurisdictions, complicating legal proceedings and necessitating international cooperation (Calzavara et al., 2015). As cyber threats grow in frequency and sophistication, traditional manual responses are becoming inadequate, especially against intelligent malware and worms that adapt dynamically to target environments (Kleinmann & Wool, 2016). In this context, artificial intelligence (AI) emerges as a promising technological enabler for enhancing cybersecurity. By leveraging advanced algorithms, adaptive learning models, and intelligent agents, AI can identify and mitigate threats in real time, potentially revolutionizing digital defense strategies (Guan & Ge, 2017; Xiao, 2018). AI's capacity to autonomously process vast datasets and detect complex patterns allows it to respond to cyber incidents with a speed and precision that far exceeds human capability, underscoring its critical role in future cyber defense architectures.

The application of AI in combating computer crimes is multifaceted, encompassing both detection and prevention mechanisms. AI's origin as a discipline in 1956 aimed at creating machines capable of solving problems requiring intelligence, which directly aligns with cybersecurity's need for systems capable of autonomously analyzing vast datasets and making informed decisions. Intelligent agents—defined as autonomous entities perceiving and acting upon their environment—utilize various techniques such as artificial neural networks, fuzzy logic, evolutionary computing, and artificial immune systems to adapt to dynamic and hostile digital landscapes. In detection, AI-driven

multi-agent systems have proven effective against worms, distributed denial-of-service (DDoS) attacks, and adaptive intrusions (Gou et al., 2006; Kotenko & Ulanov, 2007; Phillips et al., 2006). Prevention efforts, inspired by biological immune systems, have led to spam detection frameworks (Sirisanyalak & Sornil, 2007) and adaptive wireless network security models (Lebbe et al., 2007). Genetic algorithm-based intrusion detection systems have further enhanced classification accuracy, identifying anomalies before they escalate into breaches (Ojugo et al., 2012). These developments illustrate that AI can address cybercrime not merely reactively but proactively, embedding resilience into digital infrastructures.

In improving cybersecurity performance, AI's impact is particularly visible in data and network security, intrusion detection, botnet identification, email security, account protection, authentication, and application security. AI tools excel at anomaly detection by uncovering patterns in data that human operators may overlook, thereby preventing data leaks caused by unauthorized access or misconfigured systems (Kowert, 2017). Machine learning models—both supervised and unsupervised—play a central role in classifying network traffic, identifying malicious activity, and clustering suspicious behaviors (Bowman & Huang, 2019). Intrusion detection systems now integrate anomaly-based detection using AI algorithms, enabling the identification of deviations from established baselines (Ghorbanpour Vishkasouqi & Mirazimi Tabalvandani, 2023). In combating botnets, AI mitigates the challenge of distinguishing malicious traffic from legitimate flows by applying classification and clustering techniques (Bakhshi & Veisi, 2019). AI-driven spam filters employ neural networks, support vector machines, Bayesian networks, and natural language processing to counteract phishing attacks (Keshavarz & Hosseini, 2023). Additionally, platforms such as Facebook leverage deep learning models like Deep Entity Classification to detect and disable fake accounts at scale (Adekunle, 2019). These examples demonstrate AI's versatility in reinforcing cybersecurity measures across multiple vectors of attack.

Authentication and anti-spoofing have also benefited from AI's integration into cybersecurity protocols. Biometric verification methods—including facial recognition, iris scanning, fingerprint analysis, and voice recognition—are now augmented with behavioral biometrics such as typing patterns and signature dynamics (Adekunle, 2019; Chang et al., 2016). AI-enhanced multi-factor authentication systems can identify covert behaviors, detect malicious objects, and secure user access as the first line of cyber defense. However, the emergence of deepfake technology and generative adversarial networks has introduced new spoofing risks, necessitating "liveness detection" methods to ensure the authenticity of biometric data (Adekunle, 2019). In application security, AI has enabled significant advancements in vulnerability detection and secure coding practices. For example, Microsoft's AI-powered tools assist developers in identifying coding flaws with remarkable accuracy, integrating security into the early stages of the software development lifecycle (Bakhshi & Veisi, 2019). Such AI applications underline the importance of embedding intelligent, adaptive security measures directly into the core of digital systems to preempt exploitation.

Despite these advancements, AI's use in cybersecurity is constrained by several critical limitations. High-quality, large-scale datasets are essential for training AI models, but acquiring such datasets is costly, time-consuming, and often impractical due to the heterogeneity and unpredictability of cyber threats (Nunes et al., 2015). Moreover, AI decision-making processes frequently lack transparency, operating as "black boxes" whose internal logic remains opaque even to developers (Schlegel, 2019). This opacity complicates accountability, particularly in legal or regulatory contexts where explainability is essential. Adversarial inputs—such as manipulated data or pixel-level image

alterations—can deceive AI systems, producing false negatives or false positives that undermine reliability. Additionally, cybercriminals themselves can exploit AI to develop more sophisticated attacks (Bahramizadeh & Sadeghzadeh, 2022). These challenges necessitate the integration of human oversight into AI-driven systems, an approach formalized as "human-in-the-loop" architectures, which combine human judgment with machine speed and scalability to optimize decision-making in complex, high-stakes environments (Ali et al., 2017).

In conclusion, artificial intelligence represents both a transformative opportunity and a multidimensional challenge in the realm of cybersecurity. Its ability to detect, prevent, and respond to cyber threats with unprecedented speed and adaptability positions it as an essential component of modern digital defense. However, realizing its full potential requires addressing data dependency, transparency, adversarial vulnerability, and the indispensable role of human oversight. A balanced integration of AI capabilities with expert human intervention offers the most promising path toward building resilient, adaptive, and ethically accountable cybersecurity ecosystems capable of countering the evolving landscape of computer crimes.

References

Adekunle, Y. (2019). Holistic exploration of gaps vis-à-vis artificial intelligence in automated teller machine and internet banking. *International Journal of Applied Information Systems*, *2*(25), 12.

Ali, M. H., Abbas, B., Al, D., Ismail, A., & Zolkipli, M. F. (2017). A New Intrusion Detection System Based on Fast Learning Network and Particle Swarm Optimization. *IEEE Access*, *6*, 232-261. https://doi.org/10.1109/ACCESS.2018.2820092

Bahramizadeh, A. H., & Sadeghzadeh, S. (2022). Artificial intelligence and challenges in ensuring cybersecurity. Proceedings of the 16th National Conference on Computer Science and Information Technology, Mazandaran, Iran.

Bakhshi, B., & Veisi, H. (2019). End to end fingerprint verification based on convolutional neural network. 27th Iranian Conference on Electrical Engineering (ICEE), USA.

Barth, A., Rubinstein, B. I. P., Sundararajan, M., Mitchell, J. C., Song, D., & Bartlett, P. L. (2012). A Learning-Based Approach to Reactive Security. *Ieee Transactions on Dependable and Secure Computing*, *9*(4), 482-490. https://doi.org/10.1109/TDSC.2011.42

Bowman, B., & Huang, H. H. (2019). Securing malware cognitive systems against adversarial attacks. 17th IEEE International Conference on Cognitive Computing (ICCC), USA.

Brenner, S. W. (2010). *Cybercrime: Criminal Threats from Cyberspace*. Greenwood Publishing Group. https://doi.org/10.5040/9798400636554

Calzavara, S., Tolomei, G., Casini, A., Bugliesi, M., & Orlando, S. (2015). A Supervised Learning Approach to Protect Client Authentication on the Web. *Acm Transactions on the Web*, *9*(1), 1-30. https://doi.org/10.1145/2754933

Chang, C., T, E., & Obando Carbajal, L. E. (2016). Biometric authentication by keystroke dynamics for remote evaluation with one-class classification. Advances in Artificial Intelligence: 29th Canadian Conference on Artificial Intelligence, Canada.

Ghorbanpour Vishkasouqi, S., & Mirazimi Tabalvandani, S. A. (2023). A comprehensive guide to the applications of artificial intelligence in cybersecurity. Proceedings of the 11th National Conference on Novel Ideas in Humanities and Engineering, Rasht, Iran.

Gordon, S., & Ford, R. (2006). On the definition and classification of cybercrime. *Journal in Computer Virology*, *2*(1), 13-20. https://doi.org/10.1007/s11416-006-0015-z

Gou, X., Jin, W., & Zhao, D. (2006). Multi-agent system for worm detection and containment in metropolitan area networks. *Journal of Electronics*, *23*(2), 259-265. https://doi.org/10.1007/s11767-004-0111-5

Guan, Y., & Ge, X. (2017). Distributed attack detection and secure estimation of networked cyber-physical systems against false data injection attacks and jamming attacks. *IEEE Transactions on Signal and Information Processing over Networks*, *4*(1), 48-59. https://doi.org/10.1109/TSIPN.2017.2749959

Keshavarz, Z., & Hosseini, H. R. (2023). Artificial intelligence in cybersecurity: Applications, challenges, and opportunities. Proceedings of the 6th National Conference on New Technologies in Electrical, Computer, and Mechanical Engineering in Iran, Tehran, Iran.

Kleinmann, A., & Wool, A. (2016). Automatic Construction of Statechart-Based Anomaly Detection Models for MultiThreaded Industrial Control Systems. *International Publication*, *8*(4), 1-21. https://doi.org/10.1145/3011018

Kotenko, I., & Ulanov, A. (2007). Multi-Agent Framework for Simulation of Adaptive Cooperative Defense Against Internet Attacks. International Workshop on Autonomous Intelligent Systems Agents and Data Mining,

Kowert, W. (2017). The foreseeability of human-artificial intelligence interactions. *Texas Law Review*, *96*(1), 181.

Lebbe, M. A., Agbinya, J. I., Chaczko, Z., & Chiang, F. (2007). Self-Organized Classification of Dangers for Secure Wireless Mesh Networks. Australasian Telecommunication Networks and Applications Conference, Australia.

Nunes, D. S., Zhang, P., & S.S, J. (2015). A survey on human-in-the-loop applications towards an internet of all. *Ieee Communications Surveys & Tutorials*, *17*(2), 944-965. https://doi.org/10.1109/COMST.2015.2398816

Ojugo, A. A., Eboka, A. O., Okonta, O. E., Yoro, R. E., & Aghware, F. O. (2012). Genetic Algorithm Rule-Based Intrusion Detection System (GAIDS). *Journal of Emerging Trends in Computing and Information Sciences*, *3*(8), 118-119.

Phillips, L., Link, H., Smith, R., & Weiland, L. (2006). *Agent-Based Control of Distributed Infrastructure Resources*. USA Department of Energy (Sandia National Laboratories).

Regev, O. (2009). On lattices, learning with errors, random linear codes and cryptography. *Journal of the ACM*, *56*(6), 1-43. https://doi.org/10.1145/1568318.1568324

Schlegel, U. (2019). Towards a rigorous evaluation of XAI methods on time series. International Conference on Computer Vision Workshop (ICCVW), New York, USA.

Singh, K. (2017). Sparse proximity based robust fingerprint recognition. International Conference on Computing, Communication and Automation (ICCCA), England.

Sirisanyalak, B., & Sornil, O. (2007). An artificial immunity-based spam detection system. IEEE Congress on Evolutionary Computation (CEC 2007), USA.

Tsai, C. F., Hsu, Y. F., Lin, C. Y., & Lin, W. Y. (2009). Intrusion detection by machine learning: A review. *Expert Systems with Applications*, *36*, 11994-12333. https://doi.org/10.1016/j.eswa.2009.05.029

Xiao, R. (2018). Attacking network isolation in software-defined networks: New attacks and countermeasures. International Conference on Computer Communication and Networks (ICCCN), Japan.

Zhu, Y., & Tan, Y. (2011). A local-concentration-based feature extraction approach for spam filtering. *Ieee Transactions on Information Forensics and Security*, *6*(2), 486-490. https://doi.org/10.1109/TIFS.2010.2103060