هشتمین کنفرانس ملی مباحث نوین در کامپیوتر و فناوری اطلاعات
8th National Conference on Advanced Topics in Computer and Information Technology
۲۹ آذر ماه ۱۴۰۳

استفاده از الگوریتم‌های هوش مصنوعی جهت شناسایی و تشخیص انواع بدافزار در تلفن‌های همراه

الهه جان مشهدی (۱)
گروه مهندسی کامپیوتر، واحد تهران جنوب، دانشگاه آزاد اسلامی، تهران، ایران
delijan1354@gmail.com

علیرضا صفائی(۲)
گروه مهندسی کامپیوتر، واحد تهران جنوب، دانشگاه آزاد اسلامی، تهران، ایران
alisafasafaei@yahoo.com

چکیده : با پیشرفت فراوان در فناوری ارتباطات و گسترش استفاده از تلفن همراه و دستگاه‌های هوشمند، سیستم عامل اندروید به عنوان یکی از پرکاربردترین سیستم‌های عامل موبایل در جهان مطرح شده است. با افزایش استفاده از اپلیکیشن‌ها و برنامه‌های موبایلی، بدافزارها نیز به صورت گسترده‌تر و پیچیده‌تری در این سیستم‌ها نفوذ کرده‌اند. استفاده از هوش مصنوعی و یادگیری ماشین بعنوان یکی از روش‌های نوظهور و امیدوار کننده توانسته است کاربرد بسیاری جهت شناسایی و تشخیص انواع حملات و بدافزارها داشته باشد. لذا در این تحقیق بر آن شدیم تا با استفاده از هوش مصنوعی روشی ارائه دهیم که با دقت بالایی به شناسایی و تشخیص انواع بدافزار جهت حفظ محرمانگی اطلاعات شخصی کاربران در سیستم‌های رایانه‌ای و تلفن همراه بپردازد. نتایج بدست آمده از پیاده‌سازی روش پیشنهادی حاکی از آن است که دقت بدست آمده با استفاده از داده‌های آموزشی و الگوریتم گرگ خاکستری جهت تعیین معماری مناسب شبکه عصبی ۳.۱۶ درصد افزایش داشته است همچنین مقادیر معیارهای ویژگی، حساسیت، صحت و F-Measur به ترتیب ۳.۰۳ درصد، ۳.۳ درصد، ۳ درصد و ۳.۱۵ درصد افزایش داشته‌اند. همچنین مقایسه دقت روش پیشنهادی در این تحقیق با دقت برخی کارهای پیشین حاکی از آن است که روش پیشنهادی در این تحقیق نسبت به مقالات مورد مقایسه از دقت و در نتیجه عملکرد بهتری جهت شناسایی و تشخیص انواع بدافزار برخوردار است.

کلمات کلیدی: شناسایی– تشخیص -بدافزار-تلفن همراه- هوش مصنوعی

۱ - مقدمه

در دنیای امروز سیستم‌های کامپیوتری به جز ضروری زندگی ما تبدیل شده‌اند. در کنار مزیت‌های فراوان این سیستم‌ها، فرصت‌هایی برای سوء استفاده‌کنندگان به وجود آمده است. آنها با طراحی نرم افزارهای مخرب[1] که بدافزار[2] نامیده می‌شوند سعی در آسیب رساندن به سیستم‌ها و امنیت کاربران دارند. بدافزارها، گروهی از برنامه‌ها هستند که رفتارهای بدخواهانه و خواص مشابه‌ای دارند و براساس شباهت آنها را در یک خانواده قرار می‌دهند[۱]. در سال‌های اخیر یکی از مهمترین چالش‌های امنیت اطلاعات و شبکه‌های ارتباطی، افزایش روز افزون انواع بدافزارها و به دنبال آن یافتن راه‌های مناسب جهت حفاظت سیستم‌ها در مقابل آنهاست که از مهمترین دغدغه‌های برنامه‌نویسان و متخصصین امنیت اطلاعات، شناخت به موقع و یافتن راه‌های مقابله با اثرات مخرب این‌گونه

[1] Malicious sofware

[2] malware

بدافزارها می‌باشد[۲]. در مقابله با این روند، شرکت‌های ضدبدافزاری به طور جدی به دنبال راه‌های بهتری برای شناسایی بدافزار هستند. یکی از مشکلات اساسی در تشخیص بدافزارها، پیچده‌تر شدن آنها توسط تکنیک‌های مبهم‌سازی است که در این صورت مقابله با بدافزارها نیازمند تشخیص و تمایز میان بدافزارها و نرم‌افزارهای قانونی است[۳،۴]. در این راستا استفاده از هوش مصنوعی بعنوان یکی از روش‌های نوظهور و امیدوار کننده توانسته است کاربرد بسیاری جهت شناسایی و تشخیص انواع بدافزارها داشته باشد. یادگیری ماشین به عنوان یکی از شاخه‌های وسیع و پرکاربرد هوش مصنوعی به تنظیم و اکتشاف شیوه‌ها و الگوریتم‌هایی می‌پردازد که براساس آنها، سیستم‌ها توانایی یادگیری پیدا می‌کنند و قادر خواهند انواع بدافزار را شناسایی و تشخیص دهند[۵،۶].

۲ – کارهای پیشین

Banumathi و همکاران در سال ۲۰۲۴ جهت شناسایی بدافزارها استفاده از حافظه کوتاه مدت را پیشنهاد نمودند و از الگوریتم ترکیبی ملخ و جستجوی فاخته برای انتخاب ویژگی‌های موثر استفاده کردند. نتایج تجربی نشان می‌دهد که روش پیشنهادی با دقت ۹۸.۹۵ درصد از عملکرد خوبی جهت شناسایی بدافزارها برخوردار است[۷]. Santosh و همکاران در سال ۲۰۲۴ جهت تشخیص بدافزار اندروید از الگوریتم‌های بهینه‌سازی شعله پروانه، بهینه‌سازی چندوجهی و بهینه‌سازی نهنگ برای انتخاب ویژگی‌های موثر و از الگوریتم‌های درخت تصمیم، جنگل تصادفی، نزدیکترین همسایه، ساده‌ترین بیز و نزدیکترین مرکز برای طبقه‌بندی استفاده کردند. نتایج بدست آمده نشان داده است که ترکیب الگوریتم‌های بهینه‌سازی نهنگ و نزدیکترین همسایه نسبت به سایر الگوریتم‌های مورد مقایسه با دقت ۹۸.۵۳ درصد از عملکرد بهتری جهت تشخیص بدافزار اندروید برخوردار است[۸]. Kattamuri و همکاران در سال ۲۰۲۳ جهت تشخیص بدافزار و تهدید سایبری از درخت تصمیم و الگوریتم‌های بهینه‌سازی کلونی مورچه‌ها، بهینه‌سازی جستجوی فاخته و بهینه‌سازی گرگ خاکستری استفاده کردند. نتایج بدست آمده نشان داد که الگوریتم بهینه‌سازی کلونی مورچه‌ها با انتخاب ۱۲ ویژگی با دقت ۹۹.۳۷ درصد قادر به تشخیص بدافزار است[۹]. Bayat و همکاران در سال ۲۰۲۱ جهت تشخیص و بررسی حملات بدافزاری از الگوریتم‌های داده کاوی و فراابتکاری استفاده کردند. بنابراین آنها با استفاده از الگوریتم شاهین هریس به بهبود شبکه عصبی پرسپترون پرداختند. نتایج بدست آمده نشان داده است که مدل پیشنهادی با دقت ۹۷.۱۸ درصد قادر به تشخیص بدافزارها بوده و از عملکرد خوب و قابل قبولی در تشخیص انواع حملات بدافزاری برخوردار می‌باشد[۱۰]. Anggraini و همکارانش در سال ۲۰۲۰ جهت تشخیص حملات بدافزار استفاده از الگوریتم طبقه‌بندی ناوبیز و تکنیک گسسته‌سازی در فواصل 3-Interval و 5-interval را ارائه دادند. نتایج تجربی نشان می‌دهد که استفاده از الگوریتم ناوبیز جهت تشخیص بدافزار زمانی که داده‌ها از مرحله تشخیص و گسسته‌سازی عبور نکرده‌اند دارای دقتی برابر با ۶۹.۷۲ درصد بوده است، زمانی که داده‌ها با فاصله 3-Interval از مرحله تشخیص عبور کنند دقت برابر با ۷۸.۱۶ درصد خواهد بود و در زمانی که میزان فاصله برابر با 5-interval باشد به دقتی برابر با ۹۷.۷۹درصد دست یافته‌اند [۱۱].

۳ – بدافزار و انواع آن

بدافزار یا نرم‌افزار بدخواه یکی از ابزارهای اقدامات ضد امنیتی می‌باشد که بدون اجازه صاحب سیستم قصد انجام کارهای ناخواسته یا خرابکارانه را در سیستم دارد. اصطلاح بدافزار از دو واژه بدخواهانه[۳] و نرم‌افزار[۴] تشکیل شده است. از دیدگاه تخصصی بدافزار در حالت کلی برنامه‌ای کامپیوتری بوده که با هدف آسیب‌رسانی به سیستم‌های کامپیوتری طراحی شده است. بدافزارها علاوه بر اینکه با عملیات خرابکارانه به منابع و اطلاعات سیستم‌ها دسترسی پیدا می‌کنند به سرویس‌دهی و فعالیت‌های سیستم‌ها آسیب وارد کرده و با سرقت غیرمجاز اطلاعات شخصی افراد؛ سبب وارد کردن هزینه‌های گزاف به قربانیان می‌شوند. از انواع بدافزارها می‌توان به ویروس کامپیوتری[۵]، کرم[۶]، تروجان[۷]، جاسوس‌افزار[۸]، باج‌افزار[۹]، وایپر[۱۰]، بات نت[۱۱]، روت کیت[۱۲]، بدافزار ویندوز، بدافزار مک، بدافزار اندروید و .... اشاره کرد. این بدافزارها وقتی که راهی برای ورود به سیستم پیدا کنند سیستم را آلوده کرده و به سرعت به منظور یافتن قربانیان جدید شروع به اسکن مجدد کرده و خود را درون شبکه تکثیر می‌کنند. یکی از بزرگترین چالش‌ها در این زمینه مقابله با

[3] Malicious
[4] Software
[5] Computer Virus
[6] Worm
[7] Trojan
[8] Spyware
[9] Ransomware
[10] Wiper
[11] Botnet
[12] Rootkit

سرعت زیاد انتشار بدافزارها است زیرا یک سیستم آلوده به بدافزار قادر خواهد بود صدها سیستم دیگر را آلوده کند و هرکدام از این صدها سیستم دیگر نیز می‌توانند سیستم‌های دیگری را آلوده نمایند[۱۲].

۴ – روش‌های تشخیص بدافزار

تشخیص بدافزار به دو دسته کلی تشخیص مبتنی بر امضا[۱۳] و تشخیص مبتنی بر ناهنجاری[۱۴] تقسیم می‌شود. روش مبتنی بر امضا براساس بررسی کدهای مشکوک و جمع‌آوری اطلاعات به منظور توصیف اهداف مغرضانه نرم‌افزارهای مخرب است هدف اصلی این روش استخراج توالی بایت‌های ویژه از کدها به عنوان امضا است همچنین یکی دیگر از اهداف این روش جستجو برای یک امضا در داخل فایل‌های مشکوک است. بیشتر ضدبدافزارهای تجاری امروزی از مجموعه‌ای از امضاها جهت تشخیص برنامه‌های مخرب استفاده می‌کنند این کدهای مشکوک با یک توالی منحصر به فرد از ساختارهای برنامه یا بایت‌ها مقایسه می‌شود اگر این امضا در داخل پایگاه داده وجود نداشته باشد به این معناست که این فایل خوش‌خیم بوده و آلوده به بدافزار نیست بنابراین یکی از محدودیت‌هایی که در روش تشخیص مبتنی بر امضا وجود دارد نیاز به دخالت انسان در بروزرسانی پایگاه داده امضا با امضاهای جدید است همچنین گروهی از محققان نیز نشان دادند که برخی از نویسندگان بدافزارهای چندریختی قادر به شکست روش مبتنی بر امضا بوسیله تکنیک‌های مبهم‌سازی هستند بنابراین این روش تشخیص مستعد ابتلا به منفی کاذب[۱۵] است [۱۳،۱۴]. تشخیص مبتنی بر ناهنجاری عموما در دو فاز آموزش و تشخیص یا نظارت اتفاق می‌افتد. در طول فاز آموزش کشف کننده سعی در یادگیری رفتار نرمال دارد. کشف کننده در طول فاز آموزش قادر است رفتار را از میزبان یا برنامه تحت نظارت یا ترکیبی از هر دو یاد بگیرد. یکی از مزایای این روش توانایی تشخیص حملات صفر-روز می‌باشد اما از محدودیت‌های این روش می‌توان به نرخ کاذب بالای آن و پیچیدگی در تعیین اینکه کدام ویژگی باید در فاز آموزش، مورد آموزش قرار داده شود می‌باشد در واقع باتوجه به اینکه در طول فاز آموزش امکان نمایش رفتارهایی که قبلا دیده نشده است برای یک سیستم صفر نیست لذا احتمال اینکه یک روش مبتنی بر ناهنجاری مثبت کاذب را بالا ببرد نیز صفر نخواهد بود بنابراین می‌توان گفت فاز تشخیص برنامه تحت بازرسی را در زمان اجرا مورد بررسی قرار داده و آن را باتوجه به آنچه که در فاز آموزش یاد گرفته است از نظر ناپایداری مورد بررسی قرار می‌دهد به عبارتی دیگر می‌توان گفت روش تشخیصی مبتنی بر ناهنجاری جهت تصمیم‌گیری مخرب بودن یا نبودن یک برنامه تحت بررسی از دانش خود برای اینکه چه چیزی رفتار عادی را تشکیل داده است استفاده می‌کند [۱۵].

جدول ۱: مزایا و معایب روش‌های مبتنی بر امضا و ناهنجاری

تجزیه و تحلیل	مزایا	معایب
مبتنی بر ناهنجاری	زمان اسکن بسیار پایین است	ناتوان در تشخیص بدافزارهای ناشناخته
	میزان نرخ مثبت غلط پایین است	توانایی اندک در شناسایی موارد مبهم و گیج‌کننده
مبتنی بر امضا	ارائه بهترین نتایج در تشخیص بدافزارهای چندریختی	قادر به تشخیص بسیاری از ویروس‌های بسته‌بندی شده نیست

۵ – هوش مصنوعی

هوش مصنوعی شاخه‌ای از علوم کامپیوتر است که به مطالعه عامل‌های هوشمند[16] می‌پردازد. عامل هوشمند وسیله‌ای است که با درک دنیای پیرامون، احتمال دستیابی به اهدافش را بیشینه می‌کند. یک تعریف واضح‌تر از هوش مصنوعی عبارت است از توانایی یک سیستم برای تفسیر داده‌های خارجی است؛ این سیستم با دیدن داده‌ها، یاد می‌گیرد و آموخته‌های خود را برای رسیدن به اهداف خاصی به کار می‌بندد. انسان به واسطه حواس ذاتی خود اطلاعات محیطی را دریافت می‌کند و با استفاده از سیگنال‌های پیچیده مغزی تجزیه و تحلیل می‌کند که منجر به درک محیط اطراف خود می‌شود. سپس بواسطه یادگیری و تجربه، بنابر شرایط عمل تصمیم‌گیری را انجام می‌دهد. سیستم هوش ماشینی هم تقریباً مشابه همین ساز و کار، اطلاعات را دریافت می‌کند و پس از تجزیه و تحلیل با قابلیت تصمیم‌گیری خود بر اساس الگوریتم‌های تعریف شده تصمیم مقتضی را می‌گیرد. بطور کلی سیستم‌های هوشمند حجم وسیعی از داده‌های مشخص را در جهت آموزش شبکه عصبی خود و آنالیز داده‌ها استفاده می‌کند. ماشین‌های

[13] signature-based methods
[14] anomaly-based methods
[15] False negative
[16] intelligent agents

هوشمند برای پیدا کردن همبستگی و الگوها و در نهایت استفاده از این الگوها برای پیش‌بینی شرایط مشابه در آینده استفاده می‌کند. هوش مصنوعی یکی از رشته‌های علوم رایانه است که به مطالعه و توسعه دستگاه‌های هوشمند از طریق ارائه الگوریتم مناسب، می‌پردازد تا ماشین‌ها را قادر به ادراک[۱۷]، استدلال[۱۸] و یادگیری[۱۹] سازد[۱۶،۱۷].

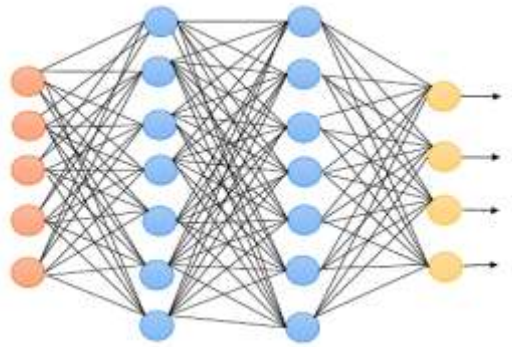
۶ –شبکه عصبی مصنوعی

در سال ۱۹۴۳ شبکه‌های عصبی مصنوعی برای اولین بار توسط مک کولاک و پیتس معرفی شدند [۱۸]. شبکه عصبی مصنوعی که از سیستم عصبی زیستی الهام گرفته شده است ایده‌ای برای پردازش اطلاعات است که همانند مغز انسان به پردازش اطلاعات می‌پردازد. امروزه به علت پشرفت تکنولوژی و پیچیده‌تر شدن مسائل در جهان استفاده از روش‌های نوین و جدید جایگزین اکثر روش‌های سنتی شده است که دیگر قادر نخواهند بود وضعیت موجود را بدرستی تخمین بزنند. از طرفی پیچیده شدن فرآیندهای موجود منجر به ایجاد مشکلاتی شده است که روش‌های سنتی قادر نخواهند بود در مورد آنها تصمیم‌گیری کنند. لذا روش‌ها و الگوریتم‌های جدیدی از قبیل شبکه عصبی مصنوعی جهت تحلیل این فرآیندها پدید آمده است. شبکه‌های عصبی مصنوعی که جهت مسائل طبقه‌بندی و شناسایی الگو طراحی شده‌اند، زیر شاخه‌ای از هوش مصنوعی هستند که قادر به درک ارتباطات و کشف نظم بین مجموعه‌ای از متغیرها می‌باشند. این شبکه از مجموعه نورون‌های عصبی با وزن‌های متفاوت با یکدیگر تشکیل شده است که پیام‌های ارسالی از هر نورون با یکدیگر جمع شده و بصورت یک پیام به مغز ارسال می‌گردد. در شبکه‌های زیستی اتصال نورون‌ها از طریق سیناپس و دندریت انجام می‌شود. دندریت یک نورون به سیناپس نورونی دیگر وصل می‌شود و سیناپس آن به دندریت یک نورون دیگر. نورون‌ها اغلب اطلاعات را از یک نورون دریافت می‌کنند و سپس آن را به نورون دیگری می‌فرستند. شبکه عصبی مصنوعی از اجزای سازنده‌ای همچون لایه‌ها و وزن‌ها تشکیل شده است به نحوی که رفتار این شبکه‌ها به ارتباط بین این اجزا وابسته می‌باشد. شبکه‌های عصبی همانطور که در شکل(۱) نشان داده شده است دارای سه نوع لایه هستند که وظیفه دریافت داده‌ها، پردازش و تولید کمیت خروجی را برعهده دارند و عبارتند از [۱۹]:

- **لایه ورودی[۲۰] :** لایه ورودی وظیفه‌ای شبیه به دندریت‌ها در شبکه عصبی زیستی دارد و تعداد نورون‌های آن با متغیرهایی که برای شبیه‌سازی یک مدل مورد نیاز است برابر می‌باشد.

- **لایه‌های میانی (مخفی یا پنهان)[۲۱] :** این لایه‌ها معادل بدنه هسته[۲۲] در شبکه عصبی زیستی هستند و تعداد نورون‌های این لایه‌ها بوسیله آزمون و خطا مشخص می‌گردد. با استفاده از ورودی‌ها و مقادیر وزن بین این لایه‌ها نحوه عملکرد لایه‌های پنهان تعیین می‌شود. فعال و غیرفعال بودن هر واحد پنهان بوسیله وزن‌های بین واحدهای ورودی و پنهان تعیین می‌شود.

- **لایه خروجی[۲۳] :** خروجی نهایی شبکه در این لایه نشان داده می‌شود. عملکرد واحد خروجی به فعالیت واحد پنهان و وزن ارتباط بین واحد پنهان و خروجی بستگی دارد.



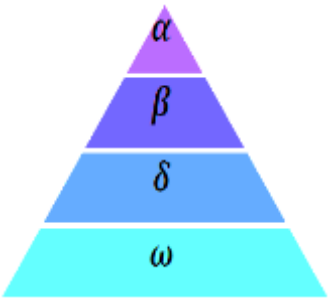
شکل ۱: یک مدل ساده از شبکه‌های عصبی مصنوعی

۷ – الگوریتم گرگ خاکستری

الگوریتم گرگ خاکستری توسط میرجلیلی و همکاران ایشان در سال ۲۰۱۴ بر مبنای شکار دسته جمعی آنها ارائه شده است[۲۰]. گرگ خاکستری از خانواده گرگ‌های کشور کانادا است. گرگ‌های خاکستری در بالای زنجیره غذایی قرار دارند و ترجیح می‌دهند به صورت گروهی زندگی کنند. بطور میانگین، گروه‌های آنها ۵-۱۲ نفر است. جالب توجه است که آنها حکومت حاکم اجتماعی خیلی سختگیرانه‌تری دارند. به این

[17] Perception
[18] Reasoning
[19] Learning
[20] Input Layer
[21] Hidden Layer
[22] Nucleus
[23] Output Layer

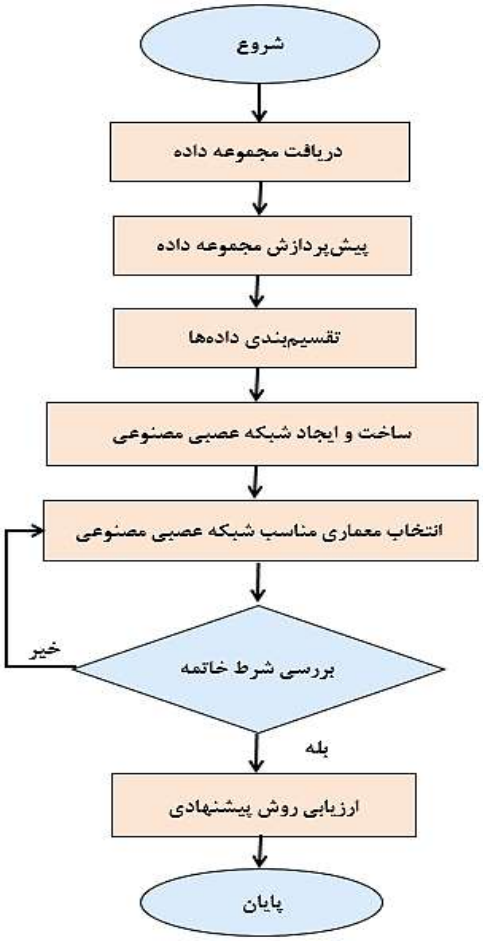
صورت که گرگ آلفا، گرگ حاکم نیز در گروه نامیده می‌شود، چرا که دستورات او باید توسط گروه پیروی شود. آلفاها اساساً مسئول تصمیم‌گیری درباره شکار، جای خواب، زمان حرکت، و مانند آن هستند. دومین سطح درجه‌بندی گرگ‌های خاکستری بتا است، بتاها گرگ‌های تحت امر آلفا هستند که به آلفا در تصمیم‌گیری و دیگر فعالیت‌های گروه کمک می‌کنند. گرگ بتا، احتمالا بهترین کاندیدا برای آلفا شدن است و نقش یک معاون را برای آلفا و ناظم را برای گروه بازی می‌کند. پایین‌ترین طبقه گرگ خاکستری امگا است. گرگ‌های امگا نقش قربانی را برای سایر اعضای گروه دارند. آنها آخرین گرگ‌هایی هستند که اجازه خوردن غذا را دارند. اگر گرگی آلفا، بتا یا امگا نباشد او فرمانبردار(یا دلتا) نامیده می‌شود. گرگ‌های دلتا از آلفاها و بتاها تبعیت می‌کنند و به امگاها حکمرانی می‌کنند. پیشاهنگ‌ها و دیده‌بان‌ها، به این گروه تعلق دارند.



شکل ۲: سلسله مراتب گرگ خاکستری

۸ –پیاده سازی روش پیشنهادی

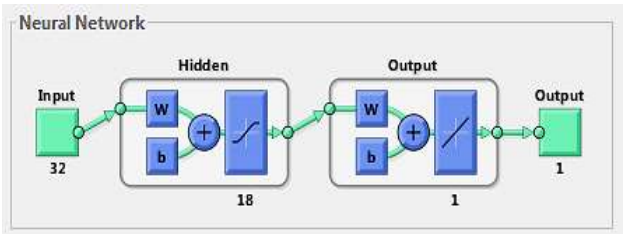
در این تحقیق بر آن شدیم تا با استفاده از هوش مصنوعی روشی ارائه دهیم که با دقت بالایی به شناسایی و تشخیص انواع بدافزار بپردازد. ایده اصلی روش پیشنهادی استفاده از الگوریتم شبکه عصبی مصنوعی برای طبقه‌بندی و استفاده از الگوریتم گرگ خاکستری جهت بهبود شبکه عصبی مصنوعی می‌باشد. فلوچارت روش پیشنهادی در شکل (۳) نشان داده شده است.



شکل ۳: فلوچارت روش پیشنهادی

مراحل روش پیشنهادی به‌طور خلاصه و طبق آنچه که در شکل (۳) نشان داده شده است بدین صورت است که پس از دریافت مجموعه داده مربوطه از سایت Kaggle[۲۴] داده‌های آن مورد پیش‌پردازش قرار گرفته و سپس داده‌ها بصورت تصادفی و با نسبت ۷۰ به ۳۰ به دو دسته داده‌های آموزشی و آزمایشی تقسیم‌بندی می‌شوند و در ادامه شبکه عصبی مصنوعی ایجاد و تعداد لایه‌ها و نورون‌های میانی آن توسط الگوریتم گرگ خاکستری تعیین و با استفاده از داده‌های آموزشی مورد آموزش قرار داده می‌شود. در پایان با استفاده از داده‌های آزمایشی به ارزیابی روش پیشنهادی پرداختیم.

[24] www.kaggle.com/nsaravana/malware-detection



شکل ۴: معماری شبکه عصبی مصنوعی

۹ – معیارهای ارزیابی روش پیشنهادی

در این تحقیق برای سنجش و ارزیابی عملکرد روش پیشنهادی براساس ماتریس درهم‌ریختگی از معیارهای دقت، حساسیت، ویژگی، صحت، F-Measure استفاده می‌شود که عبارتند از :

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (۱)$$

$$Sensitivity = \frac{TP}{TP+FN} \times 100 \quad (۲)$$

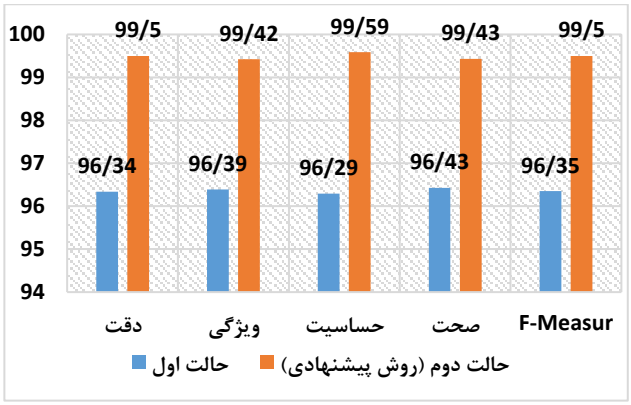
$$Specificity = \frac{TN}{FP+TN} \times 100 \quad (۳)$$

$$Precision = \frac{TP}{TP+FP} \times 100 \quad (۴)$$

$$F-Measur = 2 \times \left(\frac{sensitivity \times precision}{sensitivity+precision} \right) \quad (۵)$$

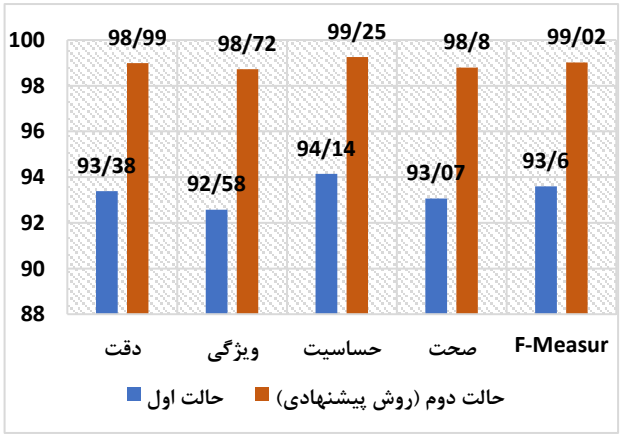
۱۰- نتایج بدست آمده

روش پیشنهادی را در این تحقیق در دو حالت مورد بررسی قرار گرفت. در حالت اول شبکه عصبی مصنوعی طراحی شده را بدون استفاده از الگوریتم گرگ خاکستری و در حالت دوم روش پیشنهادی را توسط الگوریتم گرگ خاکستری بهبود داده و به تعداد ۱۰۰۰ مرتبه مورد آموزش و ارزیابی قرار داده و اجرا کردیم. در ادامه مقایسه نتایج بدست آمده در حالت اول و حالت دوم برای داده‌های آموزشی و آزمایشی انجام در شکل‌های (۵) و (۶) نشان داده شده است.



شکل ۵: نمودار مقایسه میانگین نتایج بدست آمده در هر دو حالت برای داده های آموزشی

باتوجه به نتایج بدست آمده از پیاده‌سازی روش پیشنهادی که در شکل (۵) نشان داده شده است دقت بدست آمده جهت شناسایی و تشخیص انواع بدافزار در حالت دوم که از الگوریتم گرگ خاکستری جهت تعیین معماری مناسب شبکه عصبی پیشنهادی استفاده نمودیم نسبت به حالت اول ۳.۱۶ درصد افزایش داشته است همچنین مقادیر معیارهای ویژگی، حساسیت، صحت و F-Measur نیز در حالت دوم به ترتیب ۳.۰۳ درصد، ۳.۳ درصد، ۳ درصد و ۳.۱۵ درصد نسبت به حالت اول افزایش داشته است.



شکل ۶: نمودار مقایسه میانگین نتایج بدست آمده در هر دو حالت برای داده‌های آزمایشی

باتوجه به نتایج بدست آمده از پیاده‌سازی روش پیشنهادی که در شکل (۶) نشان داده شده است دقت بدست آمده جهت شناسایی و تشخیص

انواع بدافزار در حالت دوم که از الگوریتم گرگ خاکستری جهت تعیین معماری مناسب شبکه عصبی پیشنهادی استفاده نمودیم نسبت به حالت اول ۵.۶۱ درصد افزایش داشته است همچنین مقادیر معیارهای ویژگی،حساسیت، صحت و F-Measur نیز در حالت دوم به ترتیب ۶.۱۴ درصد، ۵.۱۱ درصد، ۵.۷۳ درصد و ۵.۴۲ درصد نسبت به حالت اول افزایش داشته است.

۱۱- نتیجه گیری

در سال‌های اخیر یکی از مهمترین چالش‌های امنیت اطلاعات و شبکه‌های ارتباطی، افزایش روز افزون انواع بدافزارها و به دنبال آن یافتن راه‌های مناسب جهت حفاظت سیستم‌ها در مقابل آنهاست که از مهمترین دغدغه‌های برنامه‌نویسان و متخصصین امنیت اطلاعات، شناخت به موقع و یافتن راه‌های مقابله با اثرات مخرب این‌گونه بدافزارها می‌باشد. در این راستا استفاده از الگوریتم‌های هوش مصنوعی بعنوان یکی از روش‌های نوظهور و امیدوار کننده توانسته است کاربرد بسیاری جهت شناسایی و تشخیص انواع بدافزارها داشته باشد که لذا در این تحقیق بر آن شدیم تا با استفاده از هوش مصنوعی روشی ارائه دهیم که با دقت بالایی به شناسایی و تشخیص انواع بدافزار بپردازد. نتایج بدست آمده از پیاده‌سازی روش پیشنهادی حاکی از آن است که دقت بدست آمده با استفاده از داده‌های آموزشی و الگوریتم گرگ خاکستری جهت تعیین معماری مناسب شبکه عصبی ۳.۱۶ درصد افزایش داشته است همچنین مقادیر معیارهای ویژگی، حساسیت، صحت و F-Measur به ترتیب ۳.۰۳ درصد، ۳.۳ درصد، ۳ درصد و ۳.۱۵ درصد افزایش داشته‌اند. همچنین مقایسه دقت روش پیشنهادی در این تحقیق با دقت برخی کارهای پیشین حاکی از آن است که روش پیشنهادی در این تحقیق نسبت به مقالات مورد مقایسه از دقت و در نتیجه عملکرد بهتری جهت شناسایی و تشخیص انواع بدافزار برخوردار است.

مراجع

[1] Alhammad, N., Alajlani, M., Abd-Alrazaq, A., Epiphaniou, G., & Arvanitis, T. (2024). Patients' Perspectives on the Data Confidentiality, Privacy, and Security of mHealth Apps: Systematic Review. *Journal of Medical Internet Research*, *26*, e50715.

[2] Almaiah, M. A., Al-Rahmi, A., Alturise, F., Hassan, L., Lutfi, A., Alrawad, M., ... & Aldhyani, T. H. (2022). Investigating the effect of perceived security, perceived trust, and information quality on mobile payment usage through near-field communication (NFC) in Saudi Arabia. *Electronics*, *11*(23), 3926.

[3] Ullah, I., Boreli, R., & Kanhere, S. S. (2023). Privacy in targeted advertising on mobile devices: a survey. *International Journal of Information Security*, *22*(3), 647-678.

[4] Phattanaviroj, T., Moslehpour, M., & Walawalkar, A. M. (2024). Data Ethics and Privacy. In *Challenges in Large Language Model Development and AI Ethics* (pp. 321-353). IGI Global.

[5] Zhang, D., Lim, J., Zhou, L., & Dahl, A. A. (2021). Breaking the data value-privacy paradox in mobile mental health systems through user-centered privacy protection: a web-based survey study. *JMIR Mental Health*, *8*(12), e31633.

[6] Khalek, S. A., Behera, C. K., & Samanta, T. (2023). An integrated framework for understanding information disclosure behaviour in mobile payment services. *Journal of Financial Services Marketing*, 1-22.

[7] Banumathi, C. S., & Rajendra, A. B. (2024). Hybridized grasshopper optimization and cuckoo search algorithm for the classification of malware. *Bulletin of Electrical Engineering and Informatics*, *13*(5), 3616-3627.

[8] Santosh Jhansi, K., Chakravarty, S., & Ravi Kiran Varma, P. (2022). A two-tier fuzzy meta-heuristic hybrid optimization for dynamic android malware detection. *SN Computer Science*, *4*(2), 117.

[9] Kattamuri, S. J., Penmatsa, R. K. V., Chakravarty, S., & Madabathula, V. S. P. (2023). Swarm optimization and machine learning applied to PE malware detection towards cyber threat intelligence. *Electronics*, *12*(2), 342.

[10] Bayat. H, Shabirinia. Misaa. (2021). Conference on Advanced Topics in Computer and Information Technology

[11] Anggraini, I., Kunang, Y. N., & Firdaus, F. (2020). Penerapan Naive Bayes Pada Detection Malware dengan Diskritisasi Variabel. Telematika, 13(1), pp.11-21

[12] Sandeep, H. R. (2019, May). Static analysis of android malware detection using deep learning. In

2019 International Conference on Intelligent Computing and Control Systems (ICCS) (pp. 841-845). IEEE.

[13] Saad, S., Briguglio, W., & Elmiligi, H. (2019). The curious case of machine learning in malware detection. *arXiv preprint arXiv:1905.07573*.

[14] Shatnawi, A. S., Yassen, Q., & Yateem, A. (2022). An android malware detection approach based on static feature analysis using machine learning algorithms. *Procedia Computer Science*, *201*, 653-658.

[15] Liu, Y., & Wang, Y. (2019, March). A robust malware detection system using deep learning on API calls. In *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)* (pp. 1456-1460). IEEE.

[16] Minh, D., Wang, H. X., Li, Y. F., & Nguyen, T. N. (2022). Explainable artificial intelligence: a comprehensive review. *Artificial Intelligence Review*, 1-66.

[17] Sun, Z., Sandoval, L., Crystal-Ornelas, R., Mousavi, S. M., Wang, J., Lin, C., ... & John, A. (2022). A review of earth artificial intelligence. *Computers & Geosciences*, *159*, 105034.

[18] McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics, 5(4), pp.115-133

[19] Zadeh, L. A. (1996). Fuzzy sets. In *Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers by LotfiA Zadeh* (pp. 394-432).

[20] Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey Wolf Optimizer Adv Eng Softw 69: 46–61.