پنجمین کنفرانس ملی مباحث نوین در کامپیوتر و فناوری اطلاعات
5th National Conference on Advanced Topics in Computer and Information Technology
۱۴ آذر ماه ۱۴۰۰

بررسی و تشخیص انواع حملات بدافزاری با استفاده از الگوریتم شاهین هریس

حمیدرضا بیات [۱]
hamidrezabayat681@gmail.com

میساء شبیری نیا[۲]
meisa_shobeiri@yahoo.com

چکیده : امروزه خطر حملات بدافزاری یکی از مهمترین چالش‌های موجود در اینترنت به شمار می‌رود از این رو پژوهشگران طی سال‌های اخیر تلاش‌های زیادی جهت شناسایی و مقابله با اینگونه حملات با استفاده از الگوریتم‌های داده‌کاوی و فراابتکاری داشته‌اند. لذا در این مقاله بر آن شدیم تا با استفاده از الگوریتم شاهین هریس و بهبود شبکه عصبی مصنوعی به شناسایی و تشخیص انواع حملات بدافزاری با دقت بالایی بپردازیم. بدین منظور برای پیاده‌سازی روش پیشنهادی در این مقاله پس از دریافت مجموعه داده موردنظر از سایت Kaggle در بازه بین صفر و یک مورد نرمال‌سازی قرار گرفت و در ادامه داده‌های موجود با نسبت ۷۰ به ۳۰ درصد به دو مجموعه داده آموزشی و مجموعه داده آزمایشی تقسیم‌بندی شده و در ادامه شبکه عصبی ایجاد و مورد آموزش قرار گرفته شد. در نهایت شبکه عصبی بهبود یافته بوسیله الگوریتم شاهین هریس مورد ارزیابی قرار گرفت و عملکرد آن را مورد بررسی قرار دادیم. نتایج بدست آمده نشان می‌دهد دقت به میزان ۱,۶۱ درصد، حساسیت ۱,۳۹ درصد ویژگی به میزان ۱,۳۷ درصد بهبود داشته است که با توجه به نتایج بدست آمده در می‌یابیم روش پیشنهادی از عملکرد خوب و قابل قبولی در تشخیص انواع حملات بدافزاری برخوردار می‌باشد.

کلمات کلیدی: بدافزار، الگوریتم شاهین هریس، داده کاوی، تشخیص.

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[۲] که بدافزار[۳] 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

[1] www.kaggle.com/nsaravana/malware-detection
[2] Malicious sofware
[3] malware

آنهاست که از مهمترین دغدغه‌های برنامه‌نویسان و متخصصین امنیت اطلاعات، شناخت به موقع و یافتن راه‌های مقابله با اثرات مخرب این‌گونه 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 الگوریتم‌های داده‌کاوی و 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] .لذا در این تحقیق بر آن شدیم تا با استفاده از الگوریتم شاهین هریس و بهبود شبکه عصبی مصنوعی به شناسایی و تشخیص انواع حملات بدافزاری با دقت بالایی بپردازیم.

۲ – کارهای پیشین

Ayesha و همکاران جهت تشخیص حملات بدافزارها در شبکه اینترنت اشیا از شبکه عصبی مصنوعی استفاده کردند. آنها همچنین از الگوریتم‌های نزدیکترین همسایگی و ناوبیز جهت مقایسه روش پیشنهادی استفاده نمودند. نتایج بدست آمده از پیاده‌سازی روش پیشنهادی نشان داد که از دقت بالایی جهت تشخیص بدافزارها برخوردار است[۷].

Samantray و همکارش جهت تشخیص بدافزار از روش‌های انتخاب ویژگی جنگل تصادفی و k-best و برای طبقه‌بندی داده‌ها از طبقه‌بندهای Naive Bayes، Logistic Regression و SVM استفاده کردند[۸].

Anggraini و همکارانش جهت تشخیص حملات بدافزار استفاده از الگوریتم طبقه‌بندی NaïveBayes و تکنیک گسسته‌سازی در فواصل 3-Interval و 5-interval را ارائه دادند [۹].

Santosh Joshi و همکاران در مقاله‌ای برای شناسایی بدافزار استفاده از طبقه‌بندی تصادفی جنگل در ساختار داده لیست پردازش را ارائه کردند[۱۰].

رحمان و همکاران از طبقه‌بندهای مختلفی همچون Random tree، Decision stump، SVM Linear Discriminant Analysis، W-J48، KNN و W-J48 graft برای تشخیص بدافزار استفاده کردند. نتایج پیاده‌سازی آنها نشان داده است که بالاترین دقت در تشخیص فایل‌های آلوده به بدافزار مربوط به الگوریتم SVM با دقت ۸۵ درصد و کمترین مقدار دقت بدست آمده در تشخیص بدافزار مربوط به الگوریتم Linear Discriminant Analysis با دقت ۴۶،۱۵ درصد است[۱۱].

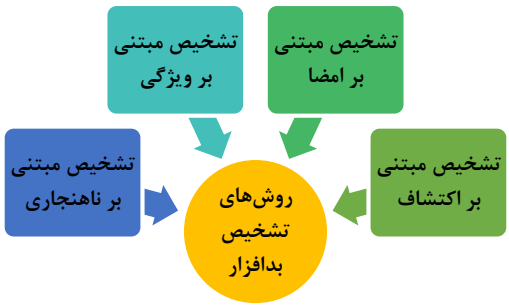
۳ – بدافزار و انواع آن

بدافزار نرم‌افزاری است که به طور خاص برای دستکاری، خرابکاری، آسیب رساندن یا دسترسی غیرمجاز به یک سیستم اطلاعاتی طراحی شده است به عبارت دیگر بدافزارها برنامه‌های رایانه‌ای هستند که به کاربر آزار رسانده یا خسارتی وارد می‌کنند و برخی از آنها تنها باعث آزار کاربر می‌شوند مثلاً وی را مجبور به انجام کاری تکراری می‌کنند یا می‌ترسانند. اما برخی دیگر به سیستم رایانه‌ای خسارت نرم‌افزاری و یا حتی سخت‌افزاری وارد کرده و داده‌ها را هدف قرار داده و یا اطلاعات آن را سرقت می‌کنند. بدافزار را گاهی آلودگی رایانه‌ای نیز خطاب می‌کنند که قادرند گوشی تلفن، تبلت و کامپیوترها را آلوده نمایند؛ بدافزار پس از ورود به سیستم می‌تواند کارهایی مانند ارسال ایمیل‌های اسپم، سرقت اطلاعات و رمزهای عبور و ... انجام دهد[۱۲]. با توجه به افزایش تهدیدات از سوی بدافزارها آشنایی با انواع آنها می‌تواند در شناخت و مقابله با آنها بسیار موثر باشد. اغلب کاربران اینترنت و رسانه‌ها از کلمه ویروس برای اشاره به هر بدافزاری که در اخبار گزارش می‌شود استفاده می‌کنند در حالی که خوشبختانه اغلب بدافزارها، ویروس نیستند و ویروس کامپیوتری فایل‌های معتبر و قانونی میزبان را به‌گونه‌ای تغییر می‌دهد تا هر زمان که فایل آلوده اجرا شد ویروس هم با آن اجرا شود. ویروس کامپیوتری خالص دیگر این روزها رایج نیست و تنها ۱۰ درصد تمام بدافزارهای موجود را تشکیل می‌دهد این اتفاق خوبی است چون ویروس تنها بدافزاری است که باقی فایل‌ها را نیز آلوده می‌کند و به

همین خاطر پاک کردن ویروس به‌تنهایی دشوار و تقریبا غیرممکن است حتی بهترین نرم‌افزارهای آنتی‌ویروس توان جدا کردن ویروس از سایر فایل‌ها را ندارند و در بیشتر موارد فایل‌های آلوده را قرنطینه یا کلا پاک می‌کنند[۱۳]. سابقه حضور کرم‌ها در سیستم‌های کامپیوتری از ویروس‌ها بیشتر است و به دوران بزرگ رایانه‌ها برمی‌گردد به نحوی که کرم‌های کامپیوتری با ظهور ایمیل در اواخر دهه ۱۹۹۰ ایجاد شدند و به مدت تقریبا ده سال کارشناسان امنیت کامپیوتر در محاصره کرم‌های مخربی بودند که به‌صورت فایل‌های پیوست در ایمیل فرستاده می‌شد[۱۴]. یکی از رایج‌ترین انواع بدافزار تروجان است که اغلب خود را به شکل ابزاری معتبر و کاربردی جا می‌زند تا کاربر را وادار به نصب خود کند، تروجان از ویروس قدیمی‌تر است اما بیشتر از هر بدافزار دیگری به کامپیوترهای کنونی آسیب زده است. اسم این بدافزار از داستان اسب تروآ گرفته شده است که در آن یونانی‌های باستان داخل اسب چوبی غول‌پیکری که به‌عنوان هدیه به شهر تروآ داده شده بود مخفی شدند و زمانی‌که اسب وارد شهر شد یونانی‌ها از آن بیرون آمدند و شهر را تصاحب کردند. بدافزار تروجان کارکرد مشابهی دارد به این صورت که مخفیانه و در قالب ابزاری کاربردی مانند بروزرسانی یا دانلود فلش وارد سیستم می‌شود و به محض ورود، حمله را آغاز می‌کند، تروجان برای دسترسی پیدا کردن به اطلاعات سیستم باید توسط کاربر اجرا شود؛ این بدافزار اغلب از طریق ایمیل یا بازدید از وب‌سایت‌های آلوده به سیستم منتقل می‌شود[۱۵]. کار جاسوس‌افزار از اسمش پیدا است جاسوسی کردن و سرک کشیدن به کامپیوتر و دستگاه‌های دیگران است؛ جاسوس‌افزار به سابقه مرورگر شما، اپلیکیشن‌هایی که استفاده می‌کنید یا پیام‌هایی که می‌فرستید دسترسی دارد. جاسوس‌افزار می‌تواند به‌صورت تروجان یا روش‌های دیگر دانلود و وارد دستگاه شود برای مثال نوار ابزاری که برای مرورگر خود دانلود می‌کنید ممکن است حاوی جاسوس‌افزاری باشد که فعالیت‌های شما را در اینترنت مشاهده می‌کند یا تبلیغات مخرب ممکن است کد جاسوس‌افزار را از طریق دانلود ناخواسته و به‌طور مخفیانه به کامپیوتر شما منتقل کنند در برخی موارد نوعی از جاسوس‌افزار به‌عنوان نرم‌افزاری با هدف کنترل استفاده از اینترنت کودک به والدین فروخته می‌شود و به گونه‌ای طراحی شده است تا نرم‌افزارهای امنیتی و آنتی‌ویروس آن را نادیده بگیرند، از طرفی برخی شرکت‌ها از جاسوس‌افزار برای نظارت مخفیانه بر کارمندان خود استفاده می‌کنند. جاسوس‌افزارها اغلب به‌آسانی قابل حذف‌اند چون برخلاف بدافزارهای دیگر قصد مخرب و شرورانه‌ای ندارند لذا کافی است فایل اجرایی جاسوس‌افزار را پیدا کنید و جلوی اجرا شدن آن را بگیرید[۱۶]. اکثر بدافزارها ترجیح می‌دهند تا جای ممکن از دید کاربر مخفی بمانند تا بتوانند اطلاعات بیشتری را دور از چشم او سرقت کنند اما باج‌افزار به خاطر ماهیت خاص خود معمولا برعکس عمل می‌کند؛ باج‌افزار اغلب از طریق فایل پیوست یا لینکی در ایمیل‌های فیشینگ وارد سیستم می‌شود آن را آلوده می‌کند و با رمزگذاری داده‌های کاربر یا بیرون انداختن او از سیستم از او باج می‌خواهد و برای اینکه به کاربر دسترسی دوباره به سیستم یا اطلاعات قفل شده‌اش را بدهد از او می‌خواهد از طریق بیت‌کوین یا رمزارزهای دیگر مبلغی به حساب هکر واریز کند[۱۷].

۴ – روش‌های تشخیص بدافزار

روش‌های تشخیص بدافزار همانطور که در شکل ۱ نشان داده شده است عبارتند از [۱۸]:



شکل ۱: روش‌های تشخیص بدافزار

تشخیص مبتنی بر ناهنجاری از دانش خود در خصوص رفتارهای عادی یک برنامه استفاده و در مورد مخرب بودن رفتارهای برنامه تحت بررسی تصمیم‌گیری می‌نماید. از مزایای این روش به توان آن در تشخیص حملات روز و حملات کاملا جدید می‌توان اشاره کرد.

روش تشخیص مبتنی بر ویژگی نوعی روش تشخیصی براساس

ناهنجاریست که تلاش برای کاهش نرخ بالای هشدارهای نادرست در روش‌های مبتنی بر ناهنجاری دارد. این تکنیک به جای تلاش برای اجرای برنامه یا سیستم، به استخراج ویژگی‌های برنامه یا سیستم می‌پردازد. این تکنیک در زمان یادگیری، به فراگیری و استخراج ویژگی‌های رفتاری عادی سیستم تحت حفاظت پرداخته و سپس این رفتارهای عادی را با رفتار برنامه‌های تحت بررسی در زمان اجرا مقایسه می‌کند.

امروزه عمومی‌ترین روش تشخیص بدافزار روش مبتنی بر امضاء است. امضاء یک خصوصیت منحصربفرد برای هر فایل است که از الگوهای استخراج شده از بدافزارهای مختلف استفاده می‌کند که این امر باعث می‌شود آنها نسبت به روش‌های دیگر تشخیص بدافزار بسیار سریع‌تر، موثرتر و کاراتر عمل کنند.

از اواخر سال ۱۹۹۰ تا ۲۰۰۸ مهمترین راه برای تشخیص بدافزار، روش تشخیصی مبتنی بر اکتشاف بود. تشخیص مبتنی بر اکتشاف شامل قوانین و الگوهای تعیین شده توسط متخصصان جهت تشخیص و شناسایی فایل‌های آلوده به بدافزار و فایل‌های بی‌خطر و سالم می‌باشد. این قوانین و الگوها برای اینکه با انواع تهدیدات مطابقت داشته و بتوانند فایل‌های بی‌خطر را به اشتباه به عنوان بدافزار شناسایی نکنند باید به اندازه کافی عمومی باشند[۱۹].

۵ –داده کاوی

داده‌کاوی به فرآیند استخراج خودکار مدل‌ها از میان انبوه داده‌ها اطلاق می‌شود. از رایج‌ترین مدل‌های داده‌کاوی طبقه‌بندی می‌باشد، که می‌تواند با استفاده از نمونه‌های از قبل طبقه‌بندی شده برای توسعه یک مدل، تعداد زیادی نمونه را طبقه‌بندی کند. فرایند طبقه‌بندی داده‌ها شامل یادگیری و طبقه‌بندی می‌باشد. در یادگیری، داده‌های آموزشی توسط الگوریتم طبقه‌بندی تجزیه و تحلیل می‌شوند اما در طبقه‌بندی، داده‌های ارزیابی برای تخمین دقت طبقه‌بندی مورد استفاده قرار می‌گیرند. اگر دقت تخمین زده شده توسط طبقه‌بندی قابل قبول باشد، می‌توان مجموعه داده‌های جدید را نیز به طبقه‌بندی اعمال کرد. برای بکارگیری تکنیک‌های داده کاوی در سیستم‌های تشخیص بدافزار، متخصصان داده کاوی روش‌های مختلفی را بکار می‌گیرند[۲۰].

۶ –شبکه عصبی مصنوعی

شبکه‌های عصبی از خانواده معماری‌های کاملا موازی الهام گرفته شده از مغز انسان می‌باشد که قادر به یادگیری و تصمیم‌گیری از نمونه‌ها است. در این مدل شبکه تجربه تولید راه حل با معنا در مسائلی که حتی زمانی که داده‌های ورودی حاوی خطا و یا نواقص هستند وجود دارد. اساسا عوامل یک شبکه عصبی مشابه سلول‌های عصبی[۴] مغز انسان بوده و از عناصر محاسباتی خیلی ساده در لایه‌ها تشکیل شده است. یک شبکه عصبی معمولی شامل لایه ورودی[۵] لایه پنهان[۶] و لایه خروجی[۷] است. هرسلول عصبی در لایه ورودی مقدار یک متغییر مستقل را نشان می‌دهد. سلول های عصبی در لایه پنهان فقط برای اهداف محاسباتی هستند و در نهایت خروجی هرکدام از سلول‌های عصبی مقدار متغییر وابسته را محاسبه می‌کند[۲۱]. شاخصی به عنوان وزن در مجموعه‌ای از اتصالات بین سلول‌های عصبی[۸] وجود دارد. این وزن‌ها، بارها در آموزش[۹] شبکه‌های عصبی برای رسیدن به راه حل مناسب تنظیم می‌شود. اتصال[۱۰] سلول‌ها باهم، رفتار شبکه عصبی را تعریف می‌نماید. هدف اصلی از یادگیری شبکه، تعیین بردار وزن است. اندازه تنظیم برای وزن یادگیری، یک تابعی از پارامترهای مجموعه شبکه است.

۷ –روش پیشنهادی

در روش پیشنهادی از شبکه عصبی مصنوعی جهت تشخیص انواع حملات بدافزاری استفاده خواهیم نمود. برای این منظور در ابتدا مجموعه داده مورد نظر را از سایت Kaggle دریافت کرده و در ادامه داده‌های موجود را پیش‌پردازش و نرمال‌سازی نمودیم. برای این منظور در این

[4] Neurons
[5] Input layer
[6] Hidden layer
[7] Output layer
[8] Nerve cell
[9] Train
[10] Connection

مقاله برای نرمالیزه کردن داده‌ها از روش نرمالیزاسیون آماری کمترین بیشترین، در بازه [1,0] طبق رابطه ۱ استفاده کردیم :

$$X_{norm} = \frac{x-\min(x)}{\max(x)-\min(x)} \qquad (۱)$$

که x مقدار داده مورد نظر جهت نرمال شدن، $\min(x)$ کمینه بردار ورودی x، $\max(x)$ بیشینه بردار ورودی x و X_{norm} مقدار نرمال شده x می‌باشد.

در گام بعدی جهت ارزیابی داده‌ها به دو مجموعه داده‌های آموزشی[۱۱] و مجموعه داده آزمایشی[۱۲] تقسیم می‌شوند؛ بدین ترتیب که بصورت تصادفی ۷۰ درصد از داده‌ها برای مجموعه آموزشی جهت آموزش روش پیشنهادی و ۳۰ درصد باقی مانده برای مجموعه آزمایشی جهت ارزیابی روش پیشنهادی انتخاب می‌شوند. پس از تقسیم‌بندی داده‌ها، شبکه عصبی ایجاد می‌شود. در ادامه پس از اینکه شبکه عصبی مصنوعی را ایجاد کردیم با استفاده از مجموعه داده آموزشی، وزن‌های تصادفی و الگوریتم پس‌انتشارخطا مورد آموزش قرار داده می‌شود. هدف از آموزش شبکه عصبی مصنوعی جستجوی بهترین اوزان شبکه عصبی مصنوعی با کمترین میانگین مربعات خطای شبکه عصبی مصنوعی است که در این مقاله قصد داریم با انتخاب بهینه اوزان و آستانه‌ها توسط الگوریتم شاهین هریس این مقدار خطا را کاهش دهیم.

الگوریتم بهینه‌سازی شاهین هریس[۱۳] در سال ۲۰۱۹ توسط حیدری و همکاران برای حل مسائل مختلف بهینه‌سازی پیشنهاد شده است. این الگوریتم برگرفته از کلونی شاهین‌های قهوه‌ای هریس است. شاهین‌ها را می‌توان یکی از باهوش‌ترین پرندگان طبیعت قرار داد. شاهین هریس به دلیل فعالیت‌های منحصر به فرد مانند همکاری با اعضای خانواده و مشارکت گروهی در شکار نسبت به شکارچیان دیگر که اغلب برای ردیابی و شکار طعمه به تنهایی عمل می‌کنند، متمایز هستند. این پرندگان اعضای خانواده خود را به خوبی می‌شناسند و بر همین اساس حرکات آنها در حین شکار با یکدیگر هماهنگ است[۲۲].

بنابراین وزن‌های شبکه عصبی باید به گونه‌ای بهینه شوند که خطای بین خروجی پیش‌بینی شده و هدف کمینه شود که در این تحقیق وزن‌های شبکه عصبی را توسط الگوریتم شاهین بهینه خواهیم نمود. بنابراین هر شاهین که عضوی از جمعیت می‌باشد یک راه حل مسئله خواهد بود که نقش یک شبکه عصبی و یا به عبارت دیگر نقش یک طبقه‌بند را دارد و بصورت برداری از مقادیر وزن‌ها و بایاس شبکه عصبی طبق رابطه ۲ می‌باشد.

$$\mathbf{(w\ ,\ b)} = [\ w_1, w_2, w_3, \ldots, w_i, b_1, b_2, \ldots, b_j\ ] \qquad (۲)$$

که در آن wi و bj به ترتیب برابر با وزن‌ها و بایاس شبکه عصبی مصنوعی می‌باشند و بردار(w,b) نیز به عنوان یک عضو از الگوریتم شاهین تعریف می‌شود.

بنابراین هر شاهین از بردار ورودی بعنوان مقادیر وزن‌های شبکه عصبی استفاده می‌کند. بر این اساس ورودی‌های آموزشی به شبکه عصبی اعمال شده و خروجی آموزشی شبکه بدست می‌آید. پس از اتمام فرایند آموزش، الگوریتم عضوی از جمعیت را که دارای بهترین مقادیر وزن‌ها است انتخاب می‌کند؛ عضو انتخاب شده در واقع عضوی است که مقدار میانگین مربعات خطای شبکه عصبی به ازای آن از سایر اعضای جمعیت کمتر می‌باشد. پس از پایان آموزش شبکه عصبی عملکرد مدل پیشنهادی با استفاده از مجموعه داده‌های آزمایشی مورد ارزیابی قرار می‌گیرد. بدین ترتیب مجموعه داده‌های آزمایشی به شبکه عصبی بهینه شده توسط الگوریتم شاهین وارد می‌شوند و خروجی مطلوب براساس معیارهای ارزیابی بیان می‌گردد.

۸ – معیارهای ارزیابی روش پیشنهادی

از جمله معیارهایی که با توجه به ماتریس آشفتگی جهت سنجش و ارزیابی کارایی روش پیشنهادی در این مقاله می‌توان استفاده نمود معیارهای دقت، حساسیت و ویژگی می‌باشند. لازم به ذکر است که در معیارهای مذکور، مفاهیم TP تعداد درستی‌های مثبت، TN تعداد

[1] Train
[1] Test

[1] Parabuteo unicinctus

درستی‌های منفی، FP تعداد خطاهای مثبت و FN تعداد خطاهای منفی در تشخیص انواع حملات نرم‌افزاری می‌باشند. طبق روابط ۳ تا ۵ استفاده می‌شود:

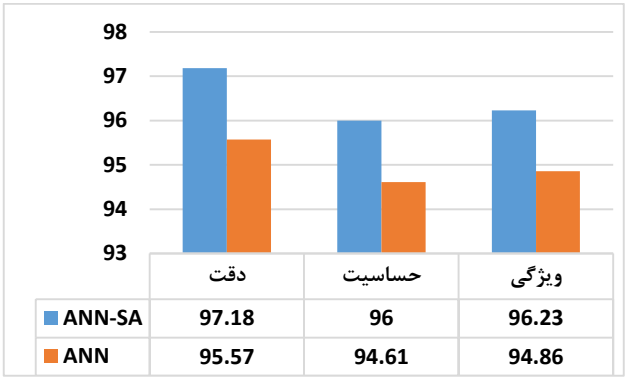
$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (۳)$$

$$Sensitivity = \frac{TP}{TP+FN} \times 100 \quad (۴)$$

$$Specificity = \frac{TN}{FP+TN} \times 100 \quad (۵)$$

۹- نتایج بدست آمده

باتوجه به معیارهای مورد بررسی نتایج بدست آمده برای داده‌های آموزشی براساس درصد در شکل ۲ مشاهده می‌شود.



شکل ۲: نتایج بدست آمده

با توجه به نتایج بدست آمده از پیاده‌سازی روش پیشنهادی که در شکل ۲ نشان داده شده است مشاهده می‌کنیم از عملکرد خوب و قابل قبولی در تشخیص انواع حملات بدافزاری برخوردار می‌باشد.

۱۰- 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 داده‌کاوی و هوش مصنوعی بعنوان یکی از روش‌های نوظهور و امیدوار کننده توانسته است کاربرد بسیاری جهت شناسایی و تشخیص انواع بدافزارها داشته باشد که در این تحقیق بر آن شدیم تا با استفاده از الگوریتم شاهین هریس و بهبود شبکه عصبی مصنوعی به شناسایی و تشخیص انواع حملات بدافزاری با دقت بالایی بپردازیم. نتایج بدست آمده نشان می‌دهد دقت به میزان ۱,۶۱ درصد، حساسیت ۱,۳۹ درصد ویژگی به میزان ۱,۳۷ درصد بهبود داشته است که با توجه به نتایج بدست آمده در می‌یابیم روش پیشنهادی از عملکرد خوب و قابل قبولی در تشخیص انواع حملات بدافزاری برخوردار می‌باشد.

مراجع

[1] Sharma, N., & Arora, B. (2020). Data mining and machine learning techniques for malware detection. In *Rising Threats in Expert Applications and Solutions* (pp. 557-567). Springer, Singapore.

[2] Yadav, P., Menon, N., Ravi, V., Vishvanathan, S., & Pham, T. D. (2020). EfficientNet Convolutional Neural Networks-based Android Malware Detection. *Computers & Security*, 102622.

[3] Komatwar, R., & Kokare, M. (2020). A survey on malware detection and classification. *Journal of Applied Security Research*, *16*(3), 390-420.

[4] Rahul, Kedia, P., Sarangi, S., & Monika. (2020). Analysis of machine learning models for malware detection. *Journal of Discrete Mathematical Sciences and Cryptography*, *23*(2), 395-407.

[5] Pastor, A., Mozo, A., Vakaruk, S., Canavese, D., López, D. R., Regano, L., ... & Lioy, A. (2020). Detection of encrypted cryptomining malware connections with machine and deep learning. *IEEE Access*, *8*, 158036-158055.

[6] Wang, C., Xu, Q., Lin, X., & Liu, S. (2019). Research on data mining of permissions

Pacific Design Automation Conference (pp. 408-413).

[16] Lysenko, S., Bobrovnikova, K., Popov, P. T., Kharchenko, V., & Medzatyi, D. (2020, June). Spyware detection technique based on reinforcement learning. In *CEUR Workshop Proceedings* (Vol. 2623, pp. 307-316).

[17] Bae, S. I., Lee, G. B., & Im, E. G. (2020). Ransomware detection using machine learning algorithms. *Concurrency and Computation: Practice and Experience*, *32*(18), e5422.

[18] Acarali, D., Rajarajan, M., Komninos, N., & Zarpelão, B. B. (2019). Modelling the spread of botnet malware in IoT-based wireless sensor networks. Security and Communication Networks..

[19] del Rey, A. M., Hernández, G., Tabernero, A. B., & Dios, A. Q. (2020). Advanced malware propagation on random complex networks. Neurocomputing, 423, 689-696.

[20] ElSawy, H., Kishk, M. A., & Alouini, M. S. (2020). Spatial firewalls: Quarantining malware epidemics in large-scale massive wireless networks. IEEE Communications Magazine, 58(9), 32-38.

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

[22] Heidari, A. A., Mirjalili, S., Faris, H., Aljarah, I., Mafarja, M., & Chen, H. (2019). Harris hawks optimization: Algorithm and applications. Future Generation Computer Systems, 97, 849–872.

mode for Android malware detection. *Cluster Computing*, *22*(6), 13337-13350.

[7] Jamal, A., Hayat, M. F., & Nasir, M. (2020). Malware Detection and Classification in IoT Network using ANN. *Mehran University Research Journal of Engineering and Technology*, *41*(1), 80-91.

[8] Samantray, O. P., & Tripathy, S. N. (2020). A Knowledge-Domain Analyser for Malware Classification. In 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA) (pp. 1-7). IEEE.

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

[10] Souri, A, Hosseini, R, 2018, A state-of-the-art survey of malware detection approaches using data mining techniques, Springer, 1-22

[11] Rehman, Z. U., Khan, S. N., Muhammad, K., Lee, J. W., Lv, Z., Baik, S. W., & Mehmood, I, "Machine learning-assisted signature and heuristic-based detection of malwares in Android devices", Computers & Electrical Engineering, vol.69, pp.828-841. 2018.

[12] Yang, L., & Liu, J. (2020). TuningMalconv: malware detection with not just raw bytes. IEEE Access, 8, 140915-140922.

[13] Komatwar, R., & Kokare, M. (2020). A survey on malware detection and classification. *Journal of Applied Security Research*, *16*(3), 390-420.

[14] Zhou, H., Hu, Y., Yang, X., Pan, H., Guo, W., & Zou, C. C. (2020). A worm detection system based on deep learning. *IEEE Access*, *8*, 205444-205454.

[15] Pan, Z., & Mishra, P. (2020). Automated test generation for hardware trojan detection using reinforcement learning. In *Proceedings of the 26th Asia and South*