ارائه روشی برای تشخیص نفوذ بر بستر اینترنت اشیا صنعتی با استفاده از یادگیری ماشین

سجاد قاسم‌زاده[۱]، رحیم اصغری[۲،*]، کوروش داداش تبار احمدی[۳]

۱- دانشجوی کارشناسی ارشد گرایش مخابرات امن و رمزنگاری، دانشگاه صنعتی مالک اشتر، تهران، ایران

۲ و ۳- استادیار مجتمع دانشگاهی برق و کامپیوتر، دانشگاه صنعتی مالک اشتر، تهران، ایران

چکیده

حفاظت از اطلاعات و جلوگیری از به سرقت رفتن داده‌ها در اینترنت اشیاء صنعتی از اهمیت بالایی برخوردار است. به‌کارگیری یک سامانه تشخیص نفوذ مناسب، شناسایی ناهنجاری‌های رخداده در شبکه را تا حد زیادی آسان می‌کند. در این مقاله، با کمک مدل‌سازی مناسب یک شبکه عصبی پیچشی که از روش‌های قدرتمند یادگیری ماشین محسوب می‌گردد، یک سامانه تشخیص نفوذ با کارایی بالا در میزان دقت در شناسایی ترافیک ناهنجار را ارائه می‌نماید. مدل پیشنهادی در محیط‌های دو کلاسه اجرا شد. هم‌چنین جهت دستیابی به نتایج مناسب، پردازش داده‌ها بر روی مجموعه داده‌های NSL – KDD پیاده‌سازی گردید که ارزیابی‌های صورت گرفته، کیفیت مناسب مدل پیشنهادشده را نشان می‌دهد.

کلمات کلیدی: اینترنت اشیاء صنعتی، امنیت شبکه، سامانه تشخیص نفوذ، یادگیری عمیق، شبکه عصبی پیچشی.

۱. مقدمه

امروزه با پیشرفت فنّاوری روزبه‌روز بر حجم و تراکم ترافیک شبکه‌های کامپیوتری افزوده می‌شود که منجر به ظهور پروتکل‌های متفاوت و جدیدی می‌شود. تجزیه‌وتحلیل این حجم زیاد داده در شبکه‌های تجاری بزرگ برای صاحبان آن شبکه‌ها مهم شده است. شرکت‌هایی که خدماتی را بر اساس تشخیص پروتکل یا تشخیص ناهنجاری ارائه می‌کنند در حال ظهور هستند و اهمیت و ارزش چنین تحلیلی را در اینترنت امروزی نشان می‌دهند[۱].

با افزایش تعداد شبکه‌های کامپیوتری، اطمینان از امنیت شبکه[1] در اولویت قرار دارد. سیستم تشخیص نفوذ سخت‌افزار یا نرم‌افزاری است که شبکه کامپیوتری را ازنظر فعالیت‌های مخرب یا نقض سیاست‌های مدیریتی یا امنیتی نظارت می‌کند و به بخش‌های مدیریت شبکه گزارش می‌دهد. روش‌های تشخیص نفوذ در شبکه به دودسته کلی تقسیم می‌شوند: تشخیص رفتار غیرعادی و تشخیص سوءاستفاده مبتنی بر امضاء[۲].

روش‌های مرسوم برای این کار، مانند روش‌های مبتنی بر نشانگر، روش‌های آماری و روش‌های پورت، در عمل دقت بالایی را نشان نمی‌دهند. علاوه بر این، این روش‌ها نیاز به تخصص دارند. هدف محققان در این زمینه این است که سامانه‌های تشخیص ترافیک و ناهنجاری‌ها نه‌تنها به‌طور خودکار و بدون دخالت عوامل خارجی عملکرد خوبی داشته باشند، بلکه از دقت تشخیصی بالایی نیز برخوردار باشند.

داده‌کاوی[2] از تکنیک‌های مختلفی مانند تکنیک‌های آماری و یادگیری ماشینی استفاده می‌کند. این تحقیق تکنیک‌های یادگیری ماشین و روش‌های طبقه‌بندی مورداستفاده برای پیاده‌سازی سامانه‌های تشخیص نفوذ را بررسی می‌کند[۳].

داده‌کاوی فرایند یافتن انواع خلاصه‌ها و مقادیر مختلف از مجموعه داده‌ها است. داده‌کاوی تکنیکی است که برای شناسایی نفوذهایی که الگوهای جدیدی را درداده های شبکه بزرگ ایجاد می‌کنند به کار می‌رود [۲]. در حال حاضر، بسیاری از

* Email:meisam.mathhome@gmail.com

[1]Network Security

[2] Data mining

محققان بر روی روش‌های تشخیص نفوذ مبتنی بر تکنیک‌های داده‌کاوی تمرکز کرده‌اند. داده‌کاوی از تکنیک‌های مختلفی مانند تکنیک‌های آماری و یادگیری ماشین استفاده می‌کند[۴].

تحقیقات انجام‌شده در سال‌های اخیر درزمینهٔ تشخیص نفوذ، تلاش کرده است تا نفوذهای شبکه را سریع‌تر و دقیق‌تر شناسایی کند؛ بنابراین، محققان از یک مجموعه داده استاندارد برای انجام آزمایش‌ها استفاده می‌کنند. استفاده از الگوریتم‌های ترکیبی، تکنیک‌های فازی، شبکه‌های عصبی، الگوریتم‌های ژنتیک و ... ازجمله روش‌هایی هستند که برای تشخیص نفوذ شبکه انجام می‌شود.

آقای مایاندی[1] و همکارانش[۵] از ترکیبی از الگوریتم خوشه‌بندی k-means برای تشخیص ناهنجاری‌ها استفاده کردند. این روش ابتدا از k-means برای بخش‌بندی نمونه‌های آزمایشی استفاده می‌کند و سپس درخت تصمیم، مرزهای تصمیم را در هر خوشه تصحیح می‌کند. آقای بوزاک[2] و همکارانش[۶] از یک رویکرد داده‌کاوی برای استخراج الگوهایی استفاده کردند که رفتار عادی را برای تشخیص نفوذ نشان می‌دهند. آن‌ها از مجموعه‌ای از قوانین اتصال فازی استخراج‌شده از داده‌های بازرسی شبکه به‌عنوان مدل‌های رفتار عادی استفاده کردند. برای تشخیص رفتار غیرعادی، آن‌ها همچنین قواعد تداعی را ترکیب کردند و شباهت آن را با مجموعه‌های استخراج‌شده از داده‌های عادی محاسبه کردند. آقای جین هو[3] و همکارانش[۳] از الگوریتم ژنتیک برای بهبود وزن و آستانه شبکه‌های عصبی استفاده کردند. آن‌ها از یک شبکه عصبی چندلایه BP برای طبقه‌بندی داده‌های عادی و حمله درزمینهٔ تشخیص نفوذ استفاده کردند. در این شبکه‌ها وزن‌ها و آستانه‌های اولیه به‌صورت تصادفی انتخاب می‌شوند. آن‌ها از یک الگوریتم ژنتیک برای انتخاب آستانه و وزن بهینه برای بهبود نتایج استفاده کردند. این روش باعث افزایش سرعت مدل شبکه عصبی BP و افزایش نرخ تشخیص نفوذ می‌شود.

باتوجه‌به اینکه در سامانه‌های تشخیص نفوذ دقت در تشخیص نفوذ شبکه‌های کامپیوتری و همچنین کاهش تعداد هشدار کاذب بسیار حائز اهمیت است، در این کار سعی شده است با مقایسه روش‌های مبتنی بر یادگیری ماشین، میزان تشخیص نفوذ افزایش یابد. هدف این مقاله، ارائه یک سامانه‌های تشخیص نفوذ با به‌کارگیری الگوریتم‌های جدید قدرتمند در یادگیری ماشین با استفاده از مجموعه داده‌های مناسب است و درنهایت، ارائه یک روش تشخیص نفوذ در یک شبکه اینترنت اشیاء صنعتی است. در بخش دوم مقاله، در خصوص مجموعه داده‌های به‌کارگیری شده در مدل پیشنهادی که از اهمیت بالایی در کسب نتایج مناسب برخوردار است، بحث می‌کنیم. در بخش سوم، شبکه عصبی پیچشی را معرفی کرده و در بخش چهارم به معرفی معیارهای ارزیابی مدل پیشنهادی می‌پردازیم. درنهایت، نتایج حاصل از پیاده‌سازی مدل بر روی مجموعه داده‌ها را ارزیابی می‌نماییم.

۲. داده مورداستفاده برای تشخیص ناهنجاری

در سال ۱۹۹۸، برنامه ارزیابی تشخیص نفوذ دارپا[4] توسط آزمایشگاه لینکلن در موسسه فناوری ماساچوست[5] توسعه و مدیریت شد. هدف آن‌ها بررسی و ارزیابی تجهیزات تشخیص نفوذ بود. آن‌ها یک مجموعه داده استاندارد حاوی تعداد زیادی ترافیک شبکه نرمال و غیر نرمال را به‌طور شبیه‌سازی‌شده بررسی کردند[۲]. مجموعه داده استاندارد برای تشخیص نفوذ به نام KDD CUP نسخه‌ای از این مجموعه است که برای مدل‌سازی در این تحقیق استفاده‌شده است[۷].

[1] Muniyandi
[2] Buczak
[3] Jian-Hu
[4] DARPA
[5] MIT

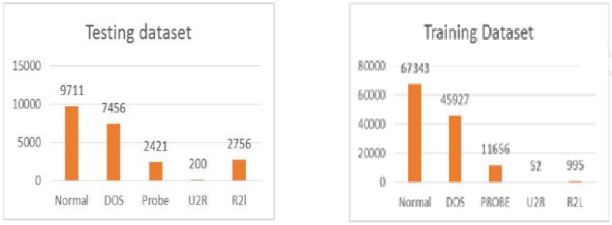
مجموعه داده KDD CUP یک نمونه شناخته‌شده در تحقیق در مورد تکنیک‌های تشخیص نفوذ است. تحقیقات زیادی برای بهبود تکنیک‌های تشخیص نفوذ درحال‌توسعه است، درحالی‌که تحقیقات روی‌داده‌های مورداستفاده برای آموزش و آزمایش مدل‌های تشخیص بسیار مهم است، زیرا کیفیت بهتر داده‌ها می‌تواند تشخیص نفوذ را بهبود بخشد[۷].

جدول ۱. ویژگی‌های داده ترافیک شبکه

نام ویژگی	شرح
hot	تعداد شاخصهای hot
num_failed_logins	تعداد تلاشهای برای ورود که با شکست مواجه شدهاند
logged_in	اگر با موفقیت وارد شده باشد مقدار ۰ در غیر اینصورت ۱
num_compromised	تعداد شرایط در معرض خطر
root_shell	اگر پوسته روت بدست آمده باشد ۰ در غیر اینصورت ۱
su_attempted	اگر دستور su root به کار گرفته شده باشد ۰ در غیر اینصورت۱
num_root	تعداد دسترسیهای root
num_file_creations	تعداد عملیاتهای ایجاد فایل
num_shells	number of shell prompts
num_access_files	تعداد عملیاتهای دسترسی کنترل فایلها
num_outbound_cmds	تعداد دستورات خارجی در یک نشست ftp
is_hot_login	اگر ورود متعلق به لیست hot باشد ۰ در غیر اینصورت ۱
is_guest_login	اگر ورود میهمان باشد ۰ در غیر اینصورت ۱
duration	طول اتصال (تعداد ثانیهها)
protocol_type	نوع پروتوکل مثل TCP , UDP و غیره
service	سرویس شبکه در مقصد برای مثال http , telnet و غیره
src_bytes	تعداد بایتهای داده از منبع به مقصد
dst_bytes	تعداد بایتهای داده از مقصد به منبع
flag	حالت نرمال یا خطای یک اتصال

آزمایشگاه لینکلن محیطی را برای دریافت داده‌های خام TCP برای یک شبکه محلی که شبیه شبکه محلی نیروی هوایی ایالات‌متحده است، پیاده‌سازی کرد. این شبکه دقیقاً مانند شبکه نیروی هوایی کار می‌کرد، اما در چندین حمله مورد آزمایش قرار گرفت. داده‌های آزمایشی خام تقریباً ۴ گیگابایت بود که از داده‌های TCP خام به‌دست‌آمده از ترافیک شبکه تشکیل‌شده است که حاصل پردازش ۵ میلیون داده مربوط به تماس در این شبکه بوده است. اتصالی که در زمان مشخصی شروع و پایان می‌یابد و جریان انتقال داده از آدرس IP مبدأ به آدرس IP مقصد طبق یک پروتکل شناخته‌شده انجام می‌شود [۸]. در مجموعه داده NSL-KDD، هر ورودی دارای ۴۳ فیلد است ۴۱ مشخصه، یک میدان رفتار بسته که رفتار یا نوع نفوذ معمولی را مشخص می‌کند و آخرین فیلد میزان دشواری در تشخیص نفوذ را نشان می‌دهد[۲]. ستون برچسب دارای ۵ دسته، یک دسته کلی و ۴ دسته نفوذ شامل DoS، U2R، R2L و Prob است. در حمله انکار سرویس، با اشباع کردن ماشین

موردنظر با درخواست‌های ارتباطی، سربار بزرگی روی سرور ایجاد می‌شود که از پاسخگویی سرور به ترافیک قانونی شبکه جلوگیری می‌کند. در ادامه برخی از این ویژگی‌ها و مقادیر داده را در جدول و شکل توضیح داده می‌شود[۴].



شکل ۱. مقادیر داده در تست و آموزش مدل

۲٫۱ استفاده از تست‌های آماری برای ویژگی‌ها

در این مقاله از فرایندهای استاندارد مانند پاک‌سازی داده‌ها و پیش‌پردازش، طبقه‌بندی، رگرسیون و حذف داده‌های پرت برای طبقه‌بندی استفاده می‌شود. در مجموعه داده NSL-KDD مجموعه داده شامل کلاس‌های مختلف حملات یعنی DoS، R2L، U2R و Prob.

۲٫۲ نرمال‌سازی داده‌ها

در مجموعه داده NSL-KDD، سه ویژگی وجود دارد که کیفی است و بقیه ویژگی‌ها کمی است. ازآنجایی‌که مقادیر ورودی باید عددی باشد، ویژگی‌های غیر عددی را به عددی تبدیل می‌کنیم. به‌عنوان‌مثال، ویژگی "protocol_type" می‌تواند سه نوع مختلف داشته باشد که شامل "tcp"، " udp " و " icmp " است.

آن‌ها را به‌صورت باینری (one hot) کدگذاری می‌کنیم شامل بردارهای باینری (۰، ۰، ۱)، (۰، ۱، ۰) و (۱، ۰، ۰) هستند. نرمال‌سازی چندین ویژگی در مجموعه داده مورداستفاده است که در آن‌ها تفاوت‌هایی بین مقادیر حداکثر و حداقل وجود دارد. ما از روش مقیاس بندی لگاریتمی برای کاهش تفاوت‌ها استفاده می‌کنیم و سپس از فرمول (۱) برای نگاشت آن‌ها به محدوده [۰, ۱] استفاده می‌کنیم[۱۰].

$$x_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (۱)$$

۳. شبکه‌های عصبی پیچشی

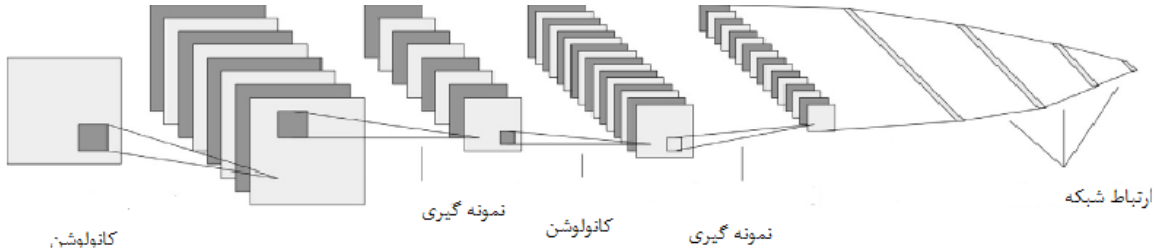
شبکه‌های عصبی پیچشی یا هم‌گشتی[1]، رده‌ای از شبکه‌های عصبی عمیق هستند که معمولاً برای انجام تحلیل‌های تصویری یا گفتاری در یادگیری ماشین استفاده می‌شوند. ساختار شبکه‌های پیچشی از فرآیندهای زیستی قشر بینایی گربه الهام گرفته‌شده است. این ساختار به‌گونه‌ای است که تک نورون‌ها تنها در یک ناحیه محدود به تحریک پاسخ می‌دهند که به آن ناحیه پذیرش گفته می‌شود. نواحی پذیرش نورون‌های مختلف به‌صورت جزئی باهم همپوشانی دارند به‌گونه‌ای که کل میدان دید را پوشش می‌دهند[۱۱].

شبکه‌های عصبی پیچشی نسبت به بقیه رویکردهای دسته‌بندی تصاویر به میزان کمتری از پیش‌پردازش استفاده می‌کنند. این امر به معنی آن است که شبکه معیارهایی را فرامی‌گیرد که در رویکردهای قبلی به‌صورت دستی فراگرفته می‌شدند[۲].

[1] CNN

این استقلال از دانش پیشین و دست‌کاری‌های انسانی در شبکه‌های عصبی پیچشی یک مزیت اساسی است. تاکنون کاربردهای مختلفی برای شبکه‌های عصبی ازجمله در بینایی کامپیوتر، سامانه‌های پیشنهاددهنده و پردازش زبان طبیعی پیشنهادشده‌اند[۱۲].

در حالت کلی، یک شبکه عصبی پیچشی، یک شبکه عصبی سلسله‌مراتبی است که لایه‌های پیچشی آن به‌صورت یک‌درمیان با لایه‌های نمونه‌گیری[1] بوده و بعد از آن‌ها تعدادی لایه تماماً متصل وجود دارد[۵].



شکل ۲. طرح کلی یک شبکه عصبی پیچشی

۳/۱ لایه‌های پیچشی

لایه‌های پیچشی یک عمل پیچش را روی ورودی اعمال می‌کنند، سپس نتیجه را به لایه بعدی می‌دهند. این پیچش درواقع پاسخ یک تک نورون را به یک تحریک دیداری شبیه‌سازی می‌کند هر نورون پیچشی داده‌ها را تنها برای ناحیه پذیرش خودش پردازش می‌کند. مشبکه کردن به شبکه‌های پیچشی این اجازه را می‌دهد که انتقال، دوران یا اعوجاج ورودی را تصحیح کنند.

اگرچه شبکه‌های عصبی پیش‌خور کاملاً همبند می‌توانند برای یادگیری ویژگی‌ها و طبقه‌بندی داده به کار روند، این معماری در کاربرد برای تصاویر به کار نمی‌رود. در این حالت حتی برای یک شبکه کم‌عمق تعداد بسیار زیادی نورون لازم است. عمل پیچش یک راه‌حل برای این شرایط است که تعداد پارامترهای آزاد را به عمیق‌تر کردن شبکه کاهش می‌دهد[۲].

۳/۲ لایه‌های ادغام

لایه‌های ادغام شبکه‌های عصبی پیچشی ممکن است شامل لایه‌های ادغام محلی یا سراسری باشند که خروجی‌های خوشه‌های نورونی در یک‌لایه را در یک تک نورون در لایه بعدی ترکیب می‌کند. به‌عنوان‌مثال روش حداکثر تجمع حداکثر مقدار بین خوشه‌های نورونی در لایه پیشین استفاده می‌کند. مثال دیگر میانگین تجمع است که از مقدار میانگین خوشه‌های نورونی در لایه پیشین استفاده می‌کند[۲].

۳/۳ وزن‌ها

شبکه‌های عصبی پیچشی وزن‌ها را در لایه‌های پیچشی به اشتراک می‌گذارند که باعث می‌شود حداقل حافظه و بیشترین کارایی به دست بیاید.

[1] Pooling

۴. معیار دقت سنجی

معیارهای ارزیابی دقت (AC) برای اندازه‌گیری عملکرد مدل استفاده می‌کنیم دقت[1] یا صحت یعنی مقدار اندازه‌گیری شده چقدر به مقدار واقعی نزدیک است[۲, ۱۰].

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (۲)$$

علاوه بر این، ما همچنین نرخ مثبت کاذب و تشخیص را حساب می‌کنیم. نرخ مثبت واقعی[2] تعداد رکوردهایی را نشان می‌دهد که دوباره به‌درستی عمل می‌کند و به‌عنوان ناهنجاری شناسایی می‌کند.

$$DR = TPR = \frac{TP}{TP + FN} \quad (۳)$$

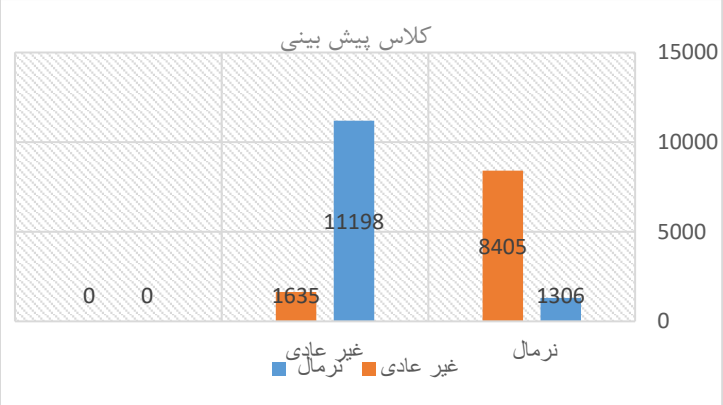
نرخ مثبت کاذب[3] اندازه‌گیری دقت یک تست است: چه یک تست تشخیصی پزشکی، یک مدل یادگیری ماشینی یا چیز دیگری باشد. در اصطلاح فنی، نرخ مثبت کاذب به‌عنوان احتمال رد کاذب فرضیه صفر تعریف می‌شود[۲, ۱۲].

$$FPR = \frac{FP}{FP + TN} \quad (۴)$$

۵. ارزیابی و تحلیل نتایج

در شبکه استفاده‌شده پیچشی درروش باینری در ابتدا داده‌ها را به دو بخش آموزش و تست طبقه‌بندی کردیم. مقدار آزمودن را برای صحت سنجی روش استفاده می‌کنیم. در مدل منفرد شبکه‌های عصبی پیچشی، در طبقه‌بندی باینری، از ۴۱ ویژگی مشخص‌شده استفاده کردیم. در شبکه‌های عصبی پیچشی، تعداد لایه‌های پنهانی که گرفته شد و تعداد دوره‌ها[4] ۱۰۰ است. در لایه اول تعداد گره را مشخص کردیم نتایج آزمایش با استفاده از سامانه رایگان Google Collab انجام شد. درنهایت مدل را آموزش داده و از تابع فعال‌ساز رلو[5] استفاده کردیم. برای تقسیم‌بندی آزمایش پنج طبقه‌بندی شامل Normal، DoS، R2L، U2R به شکل حمله تعریف شد و بقیه داده‌ها که در کلاس طبقه‌بندی بودند به‌صورت عادی شناسایی شدند. در شکل ۳، مقادیر ویژگی‌های جدول تست آمده است.

[1] Accuracy
[2] TPR
[3] FPR
[4] Epocs
[5] Relu



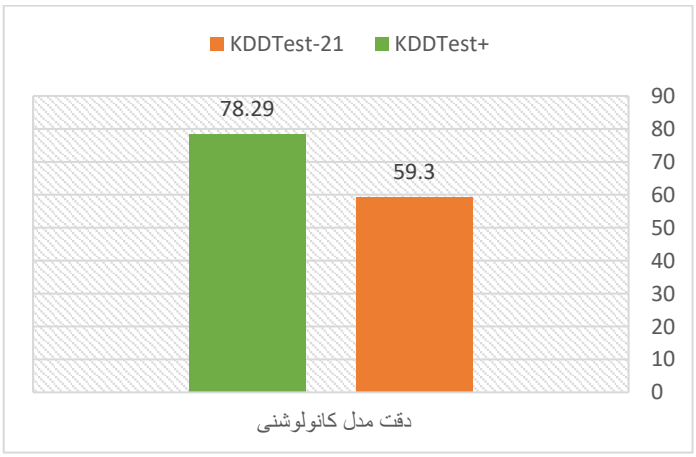
شکل ۳. تعداد داده نرمال و غیرعادی

در جدول ۲، معیارهای دقت سنجی به تفکیک آمده است.

جدول ۲. مقادیر دقت سنجی

	معیارهای ارزیابی دقت	نرخ مثبت واقعی TPR	نرخ مثبت کاذب FPR
KDDTest+	۷۸/۲۹	۰/۶	۰/۸
KDDTest-21	۵۹/۳	۰/۴	۰/۶

در شکل ۴، مقادیر دقت مدل با استفاده از روش پیچشی ترسیم‌شده است.



شکل ۴. دقت مدل پیچشی

۶. نتیجه‌گیری

تشخیص نفوذ یک ابزار مهم و کاربردی جهت برقراری امنیت در سامانه‌های کامپیوتری است. سامانه‌های تشخیص نفوذ سعی دارند نفوذهای غیرمجاز به شبکه را با توجه الگوریتم‌های خاص تشخیص دهند و می‌توان آن‌ها را به دودسته تشخیص سوءاستفاده و رفتار غیرعادی تقسیم کرد. در این مقاله، یک سیستم تشخیص نفوذ مبتنی بر دسته‌بندی دو کلاسه برای بهبود کیفیت تشخیص ناهنجاری در مقایسه با سامانه‌های موجود توسعه داده شد. در قسمت پیش‌پردازش داده با استفاده

از انتخاب ویژگی عمل کردیم و ارزیابی نتایج را انجام دادیم. در مقایسه با روش‌های سنتی، نتایج ارزیابی‌ها دقت مناسبی را نشان می‌دهد که توسط مدل پیشنهادی مبتنی بر شبکه عصبی پیچشی افزایش پیدا کرد. نتایج ارزیابی با سامانه‌های موجود نشان می‌دهد که سیستم پیشنهادی دارای دقت بیشتر است که در مدل +KDDTest مقدار دقت ۷۸/۲۹ و در -KDDTest 21 مقدار دقت به ۵۹/۳ رسیده است.

۷. مراجع

[1]. Revathi, S. and A. Malathi, A detailed analysis on NSL-KDD dataset using various machine learning techniques for intrusion detection. International Journal of Engineering Research & Technology (IJERT), 2013. 2(12): p. 1848-1853.
[2]. Li, Y. et al. Robust detection for network intrusion of industrial IoT based on multi-CNN fusion. Measurement, 2020. 154: p. 107450.
[3]. Jiang, J. et al. A dynamic ensemble algorithm for anomaly detection in IoT imbalanced data streams. Computer Communications, 2022. 194: p. 250-257.
[4]. Pei, J. et al. Personalized federated learning framework for network traffic anomaly detection. Computer Networks, 2022. 209: p. 108906.
[5]. Muniyandi, A.P. R. Rajeswari, and R. Rajaram, Network anomaly detection by cascading k-Means clustering and C4. 5 decision tree algorithm. Procedia Engineering, 2012. 30: p. 174-182.
[6]. Buczak, A.L. and E. Guven, A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications surveys & tutorials, 2015. 18(۲): p. 1153-1176.
[7]. Sharma, B. L. Sharma, and C. Lal. Anomaly detection techniques using deep learning in IoT: a survey. in 2019 International conference on computational intelligence and knowledge economy (ICCIKE). 2019. IEEE.
[8]. Deshmukh, D.H. T. Ghorpade, and P. Padiya. Improving classification using preprocessing and machine learning algorithms on NSL-KDD dataset. in 2015 International Conference on Communication, Information & Computing Technology (ICCICT). 2015. IEEE.
[9]. Hasan, M. et al. Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches. Internet of Things, 2019. 7: p. 100059.
[10]. Ingre, B. A. Yadav, and A.K. Soni. Decision tree based intrusion detection system for NSL-KDD dataset. in International conference on information and communication technology for intelligent systems. 2017. Springer.
[11]. Masoodi, F. Machine Learning for Classification analysis of Intrusion Detection on NSL-KDD Dataset. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 2021. 12(10): p. 2286-2293.
[12]. Su, T. et al. BAT: Deep learning methods on network intrusion detection using NSL-KDD dataset. IEEE Access, 2020. 8: p. 29575-29585