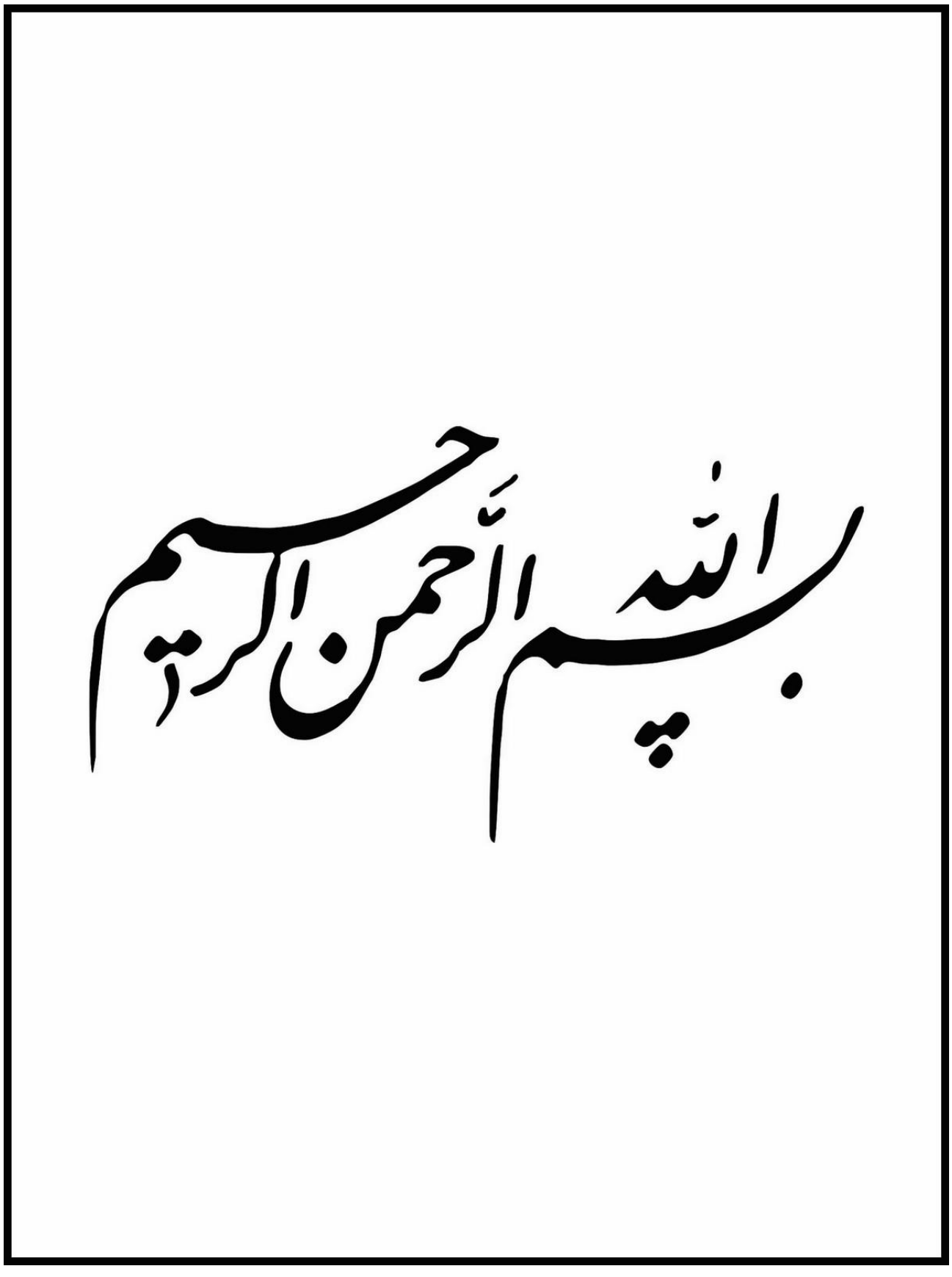

بسم الله الرحمن الرحیم

موسسه آموزش و عالی سلمان
دانشکده مهندسی

پایان نامه کارشناسی ارشد رشته مهندسی کامپیوتر گرایش شبکه های کامپوتری

بهبود دقت تشخیص نفوذ در شبکه به کمک یادگیری عمیق با الگوریتم بهینه ADAM و تابع فعالساز LEAKY RELU

نگارش
مریم مهدوی

استاد راهنما
آقای دکتر سید رضا کامل طباخ فریضنی

استاد مشاور

شهریور۱۴۰۱

(برگه اصالت یا مالکیت پایان نامه)

اینجانب مریم مهدوی دانش آموخته کارشناسی ارشد مهندسی کامپیوتر دانشکده کامپیوتر موسسه آموزش عالی سلمان پدیدآور پایان‌نامه با عنوان بهبود دقت تشخیص نفوذ به کمک یادگیری عمیق با الگوریتم بهینه ADAM و تابع فعالساز LEAKY RELU با راهنمایی جناب آقای دکتر سید رضا کامل طباخ فریضنی گواهی و تعهد می دهم که بر پایه قوانین و مقررات، از جمله بررسی تخلفات پژوهشی و همچنین مصادیق تخلفات پژوهشی، مصوب وزارت علوم تحقیقات و فناوری در تاریخ ۲۵ اسفند ۹۳:

- ☐ پایان‌نامه دستاورد پژوهش اینجانب و محتوای آن از درستی و اصالت برخوردار است.
- ☐ حقوق معنوی همه کسانی که در بدست آمدن نتایج اصلی پایان‌نامه تاثیرگذار بوده‌اند، رعایت کرده‌ام و هنگام کاربرد دستاورد پژوهش‌های دیگران در آن ، با دقت و درستی به آن استناد کرده‌ام.
- ☐ این پایان‌نامه و محتوای آن را تاکنون اینجاب یا کس دیگری برای دریافت هیچگونه مدرک یا امتیازی در هیچ‌جا ارائه نکرده‌ایم.
- ☐ همه حقوق معنوی این پایان‌نامه از آن موسسه آموزش و عالی سلمان است و آثار برگرفته از آن با وابستگی سازمانی موسسه سلمان منتشر خواهد شد.
- ☐ در همه گام‌های انجام این پایان‌نامه هرگاه به اطلاعات شخصی افراد یا اطلاعات سازمان‌ها دسترسی داشته یا آنها را به کار برده‌ام، رازداری و اخلاق پژوهش را رعایت کرده‌ام.

تاریخ امضا

برگه تایید هیئت داوران /صورت جلسه دفاع

تقدیم

این پایان نامه را با نهایت سپاس و تشکر پیشکش می کنم

به بارگاه مقدس امام هشتم امام رضا(ع) که همواره از برکات وجود ایشان بی نصیب نمانده‌ام.

وتقدیم میکنم

به دریای بی کران فداکاری وپشتیبانی

همسرم

وتقدیم می کنم

به وجود پر صلابت که همواره راهنما و حامی من بوده‌اند

پدرم

وتقدیم می کنم

به مهربان تر از جانم که همیشه مهربانی و دعای خالصانه‌اش شامل حال من شده‌است

مادرم

وتقدیم می کنم

به زیباترین هدیه زندگیم، عطر خوش لحظاتم و فرشته ی رویاهای من که از بدو شروع این راه در وجود من ریشه زد و همراه، دلگرمی من بود

دخترم

تشکر و قدردانی

حال که به لطف و عنایت حق تعالی مراحل این پایان‌نامه رو به اتمام نهاده بر خود لازم می دانم از همه کسانی که در پیشبرد اهدافم من را یاری نموده اند سپاسگزاری و قدردانی کنم

از استاد راهنمای گرانقدرم جناب آقای دکتر سید رضا کامل که وجودشان قوت قلبم بوده و همواره از دانش و راهنمایی ها و حمایت های دریغشان نهایت تشکر را دارم

و از سرکار خانم دکتر توتونیان که همواره مساعدت های لازم را بابنده داشته اند

و از جناب آقای مهندس الیاس یاری و جناب آقای مهندس سید علی کلامی نهایت تشکر بابت تجربیاتشان که بی منت در اختیار اینجانب قرار دادند

و از کتابخانه آستان قدس رضوی مشهد و دانشگاه آزاد اسلامی واحد مشهد که امکانات لازم را در اختیار اینجانب قرار دادند.

چکیده

یکی از مسائل مهم با توجه به استفاده گسترده از شبکه‌های کامپیوتری، حملات سایبری و بحث امنیت در این شبکه‌ها و پایگاه داده‌ها است. نفوذ به شبکه‌های کامپیوتری با انگیزه‌های مختلف از جمله سیاسی، نظامی، مالی و یا نشان دادن سستی و ضعف امنیتی در برنامه‌های موجود می‌باشد. از اینرو، تکنیک‌های متداول با توجه به ویژگی های مخرب جدید که به طور تصاعدی درحال افزایش می‌باشد بی اثر شده و روش‌های سنتی قادر به حفظ امنیت نمی‌باشد. در نتیجه سیستم‌های تشخیص نفوذ عهده دار شناسایی تهدیدات و حملات از سوی هر دو دسته کاربران داخلی و خارجی و ابزاری جهت تامین امنیت در سیستم های اطلاعاتی می باشند که هدف اصلی آنها جلوگیری از حمله نیست، بلکه شناسایی حملات و یادگیری الگوی رفتاری آنها از وظایف این سیستم می‌باشد. یکی از موضوعات پرچالش در این سیستم‌ها "دقت" می‌باشد که توجه محققان را در این زمینه جلب کرده است و با بهبود هرچه بیشتر "دقت" در این سیستم‌ها کارایی نیز افزایش می‌یابد.

جهت بهبود "دقت" سیستم‌های تشخیص نفوذ روش‌های متعددی ارائه شده است اما همچنان این زمینه نیاز به پژوهش و مطالعه می‌باشد. در این پایان‌نامه روشی جهت حل مسأله "دقت" از طریق کاهش ابعاد مجموعه داده NSL-KDD با تحلیل مولفه‌های اساسی(PCA) که به طور موثری در بهبود سرعت پردازش داده تاثیر می گذارد و استفاده از شبکه‌های پروسپترن چندلایه که یکی از روش‌های یادگیری عمیق (DL) می‌باشد با افزایش تعداد لایه های پنهان، از یک لایه [۱] به دو لایه پنهان که معماری این پروسپترون چهار لایه (MLP) شامل یک لایه ورودی و دو لایه پنهان و یک لایه خروجی می‌باشد سبب یادگیری کارآمدتر در این شبکه‌ها می‌شود و با الگوریتم بهینه ADAM و انتخاب تابع فعالساز LEAKY RELU در این شبکه سعی در بهبود دقت تشخیص نفوذ داریم. روش ارائه شده در شبیه‌ساز Python اجرا شده است و با آخرین کارهای انجام شده مقایسه شده است. روش پیشنهادی نشان می‌دهد دقت برای شناسایی ۵ کلاس در مجموعه داده NSL-KDD به۹۳ درصد افزایش یافته است که کارایی بیشتر روش پیشنهادی را در مقایسه با کار انجام شده نشان می‌دهد .

کلیدواژه: سیستم تشخیص نفوذ(IDS)،یادگیری عمیق(DL) پروسپترون چندلایه(MLP)، الگوریتم ADAM، تحلیل مولفه‌های اساسی(PCA)، مجموعه دادهNSL-KDD، دقت

فهرست نوشتار

۱.کلیات تحقیق.....۱۵
۱-۱ مقدمه.....۱۵
۱-۲ مسأله پژوهش.....۱۶
۱-۳ اهمیت و هدف و جنبه های نوآوری پژوهش.....۱۶
۱-۴ فرضیه های پژوهش.....۱۷
۱-۵ دامنه تحقیق.....۱۷
۱-۵ مختصری بر روش تحقیق.....۱۹
۱-۶ ساختار پایان نامه.....۱۹
۲.پیشینه تحقیق و مروری بر کارهای انجام شده.....۱۹
۲-۱ مقدمه.....۲۰
۲-۲مروری برسیستم تشخیص نفوذ.....۲۰
۲-۲-۱ تشخیص نفوذ.....۲۱
۲-۲-۲ معماری سیستم‌های تشخیص نفوذ.....۲۱
۲-۲-۲-۱ سیستم های تشخیص نفوذ مبتنی بر میزبان(HIDS).....۲۱
۲-۲-۲-۲ سیستم‌های تشخیص نفوذ مبتنی برشبکه(NIDS).....۲۲
۲-۲-۲-۳ سیستم‌های تشخیص نفوذ توزیع شده(DIDS).....۲۲
۲-۲-۳ روش‌های تشخیص نفوذ.....۲۲
۲-۲-۳-۱ روش تشخیص نفوذ مبتنی بر سوءاستفاده.....۲۲
۲-۲-۳-۲ روش تشخیص نفوذ مبتنی بر ناهنجاری.....۲۳
۲-۳ مروری بر روش‌های و اصطلاحات رایج شبکه‌های عصبی جهت تشخیص نفوذ.....۲۳
۲-۳-۱ هوش مصنوعی.....۲۴
۲-۳-۲ یادگیری ماشین.....۲۴
۲-۳-۳ یادگیری عمیق.....۲۴
۲-۳-۴ شبکه‌های مصنوعی.....۲۵
۲-۳-۴-۱ مقدمات شبکه‌های عصبی.....۲۵
۲-۳-۴-۲ نرون.....۲۵
۲-۳-۴-۳ وزن.....۲۶
۲-۳-۴-۴ بایاس.....۲۶

۲-۳-۴-۵ توابع فعالساز ۲۶
۲-۳-۴-۶ لایه‌های ورودی و خروجی و مخفی ۲۷
۲-۳-۴-۷ مدل‌های شبکه عصبی ۲۷
۲-۳-۴-۸ شبکه‌های پروسپترون چندلایه ۲۷
۲-۳-۴-۹ روش یادگیری ۲۷
۲-۳-۴-۹-۱ یادگیری کاهشی گرادیان ۲۸
۲-۳-۴-۱۰ انتشار رو به جلو ۲۸
۲-۳-۴-۱۱ انتشار به عقب ۲۸
۲-۳-۴-۱۲ تابع هزینه ۲۸
۲-۳-۴-۱۳معیار عملکرد ۲۹
۲-۳-۴-۱۴ بهینه‌سازی وزن‌ها ۲۹
۲-۳-۴-۱۵ نرخ یادگیری ۲۹
۲-۳-۴-۱۶ بسته‌ها ۲۹
۲-۳-۴-۱۷ دوره‌ها ۲۹
۲-۳-۵ چالش های آموزش در شبکه‌های عمیق ۲۹
۲-۳-۵-۱ محو گرادیان و انفجار گرادیان ۳۰
۲-۳-۵-۲ شناسایی محو گرادیان و انفجار گرادیان ۳۰
۳-۲-۶ تفاوت پارامترها و ابرپارامترها ۳۰
۲-۴ کاهش ویژگی ۳۱
۲-۵ مروری بر کارهای انجام شده ۳۱
۲-۶ جمع‌بندی ۳۶
۳. روش پیشنهادی ۳۷
۳-۱ مقدمه ۳۷
۳-۲ الگوریتم روش پیشنهادی ۳۷
۳-۳ جزئیات روش پیشنهادی ۳۸
۳-۳-۱ گام اول ۴۰
۳-۳-۲ گام دوم ۴۰
۳-۳-۳ گام سوم ۴۴
۳-۴ جمع‌بندی ۴۴
٤.روش تحلیل و ارزیابی ۴۶

۴-۱ مقدمه ٤٦
۴-۲ روش ارزیابی ٤٦
۴-۳ تحلیل ارزیابی ٤٩
۴-۳-۱ خطا(Loos) ٤٩
۴-۳-۲ صحت(Accuracy): ٥٠
۴-۳-۳ یادآوری(Recall) ٥١
۴-۳-۴ دقت (Precisions) ٥٢
۴-۳-۶ گزارش طبقه‌بندی ٥٣
۴-۴ مقایسه دقت ٥٥
۴-۴نتیجه‌گیری ٥٥
۵. نتیجه‌گیری و پیشنهاد کار آینده ٥٦
۵-۱ نتیجه‌گیری و جمع‌بندی ٥٦
۵-۲ پیشنهاد کار آینده ٥٦
٦. منابع ٥٧

فهرست کوته نوشت‌ها

IDS(Intrusion Detection System)
MLP(Multi Layer Perceptron)
FSL(Few Shot Learning)
RNN(Recurrent Neural Network)
RF(Random Forest)
NB(Naive Bayes)
AE(Auto Encoder)
SCM(Support Vector Machine)
ML (Machine Learning)
DL(Deep Learning)
CART(Classification And Regression Tree)
GDM/AG(Gradient Descent Momentum /Adaptive Gain)
GDM(Gradient Descent with Momentum)
HIDS(Host intrusion detection systems)
NIDS(Network intrusion detection systems)
DIDS(Distributed intrusion detection systems)

فهرست جداول

جدول ۲- ۱. مقایسه الگوریتم های مرور شده ۳۶
جدول ٤- ۱. توزیع انواع حملات در مجموعه داده NSL-KDD [۳۳] ٤٦
جدول ٤- ۲. تنظیم ابرپارامترها در شبکه mlp ٤۹
جدول ٤- ۳. گزارش طبقه بندی برای مجموعه داده آموزش ٥٤
جدول ٤- ٤. گزارش طبقه بندی برای مجموعه داده اعتبارسنجی ٥٤
جدول ٤- ٥. گزارش طبقه بندی برای مجموعه داده آزمایش ٥٤
جدول ٤- ٦. مقایسه دقت الگوریتم pca-dl و الگوریتم پیشنهادی ٥٥

فهرست شکل

شکل ۱- ۱. سیستم تشخیص نفوذ.....۱۵
شکل ۱- ۲. دامنه تحقیق.....۱۸
شکل ۱- ۳. ساختار کلی پایان نامه.....۱۹
شکل ۲- ۱.ارتباط بین هوش مصنوعی و یادکیری ماشین و یادگیری عمیق و شبکه‌های عصبی و شبکه‌های پروسپترون چندلایهmlp.....۲٤
شکل ۲- ۲. نمایی از یک شبکه عصبی واقعی [١٤].....۲۵
شکل ۲- ۳. معماری یک شبکه پروسپترون تمام متصل [١٤].....۲۷
شکل ۳- ۱. نمای کلی الگوریتم پیشنهادی.....۳۸
شکل ۳- ۲. فلوچارت الگوریتم پیشنهادی.....۳۹
شکل ۳- ۳. تابع فعالساز یکسوساز خطی رخنه دار [١٤].....٤۲
شکل ٤- ۱. نمونه رکورد از مجموعه داده NSL-KDD [۳٤].....٤۷
شکل ٤- ۲. نمودار همبستگی بین ویژگی های مجموعه داده.....٤۷

فهرست نمودار

نمودار٤- ١. خطا در طول آموزش مدل ٥٠
نمودار٤- ٢. صحت در طول آموزش مدل ٥١
نمودار٤- ٣. معیار یادآوری در طول آموزش مدل ٥٢
نمودار٤- ٤. نمودار دقت در طول آموزش مدل ٥٣

۱.کلیات تحقیق

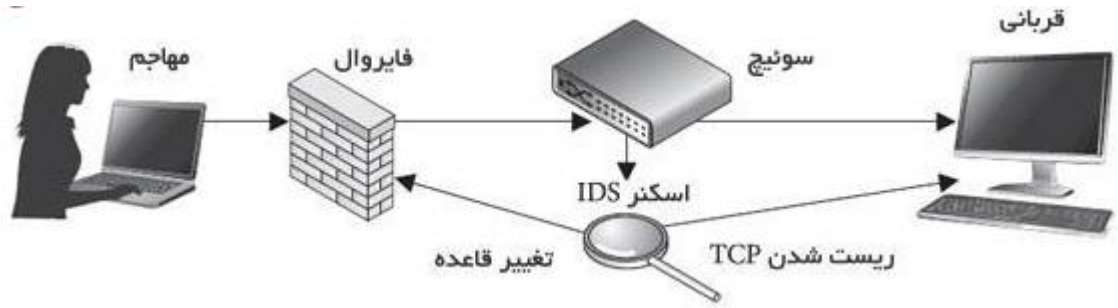
۱-۱ مقدمه

حفظ امنیت اطلاعاتی و حفظ کارایی در شبکه‌های کامپیوتری، یکی از ضروری‌ترین مسائل در سیستم‌های مخابراتی و شبکه‌های کامپیوتری می‌باشد[۲]از گذشته روش‌هایی مانند احراز هویت کاربر، رمزگذاری داده‌ها، دیواره اتش و غیره جهت برقراری امنیت در شبکه‌های کامپیوتری مورد استفاده بوده است. سیستم‌های تشخیص نفوذ به عنوان یکی از عناصر اصلی زیرساخت‌های امنیت در بسیاری از سازمان‌ها می‌باشد. این سیستم‌ها با استفاده از روش‌های تحلیلی به کشف حملات می‌پردازندف منابع حمله راشناسایی می‌کنند و هشدارهای لازم را برای مدیران شبکه ارسال می‌کنند.[۳]

از آنجا که طراحی سیستم‌های کامپیوتری به شکل ایده‌آل امکانپذیر نیست، باید تکنیک‌های حفظ اطلاعات در شبکه به کار گرفته شود. در این راستا تشخیص نفوذ در سیستم‌های کامپیوتری دارای اهمیت ویژه‌ای است.[۲]

نفوذ مجموعه اقدامات غیر قانونی است که صحت و محرمانگی به یک منبع را به خطر می‌اندازد. تشخیص نفوذ عبارت است از فرآیند شناسایی و پاسخ به فعالیت‌های مخرب که به شکل هدفمند منابع شبکه را مورد حمله قرار می‌دهد از این رو وجود سیستم‌های تشخیص نفوذ کارآمدبرای تامین امنیت شبکه از اهمیت بسیار زیادی برخوردار است . نفوذ می تواند به دو دسته داخلی و خارجی تقسیم شود. نفودهای خارجی به آن دسته نفوذهایی گفته می‌شود که توسط افراد مجاز و یا غیرمجاز از خارج از شبکه به درون شبکه داخلی صورت می گیرد و نفوذهای داخلی توسط افراد مجاز در سیستم ئ شبکه داخلی از درون خود شبکه انجام می‌پذیرد.[۲]

یکی از مزایای ایم سیستم‌ها توانایی مستند کردن نفود یا تهدید تشخیص داده شده است در نتیجه جدیدترین الگوهای حمله برای عموم قابل شناسایی و پیشگیری خواهد بود.در واقع هدف سیستم‌های تشخیص نفود جلوگیری از حمله نمی باشد بلکه هدف کشف و شناسایی حملات و تشخیص اشکالات امنیتی در شبکه‌های کامپیوتری و اعلام آن به مدید سیستم است.[۳]



شکل ۱- ۱. سیستم تشخیص نفوذ

در سیستم‌های تشخیص نفوذ معیار دقت که مربوط به شناسایی الگوی رفتار های نرمال و حمله است یکی از چالش‌های حیاتی مطرح در این حوزه می‌باشد روش‌های بسیار زیادی از جمله روش‌های یادگیری عمیق، روش های ترکیبی برای بهبود این معیار ارائه شده است.

۱-۲ مسأله پژوهش

نفوذ در شبکه‌های کامپیوتری با هدف‌های متفاوتی از جمله سیاسی، مالی، نظامی و یا نمایان کردن سستی‌های امنیتی موجود در برنامه‌های کاربردی صورت می‌گیرد. مقصود اصلی سیستم‌های تشخیص نفوذ پیشگیری از حمله نمی‌باشد بلکه شناسایی و کشف حملات و تشخیص مشکلات امنیتی است.

داده‌کاوی و یادگیری ماشین از راه‌های حیاتی یافتن علم سودمند اطلاعات در داده‌ها می‌باشد که در تحلیل انواع ترافیک شبکه کاربرد دارد چندین الگوریتم داده‌کاوی مانند منطق فازی، شبکه‌های عصبی، دستگاه بردار پشتیبان، k نزدیکترین همسایه و غیره در زمینه تشخیص نفوذ شبکه‌های کامپیوتری به کار رفته است[۴].

مهم‌ترین مساله و چالشی که در این روش‌ها وجود دارد جداسازی فعالیت‌های نرمال از حملات است که باتوجه به روش‌های ارایه شده دارای نرخ هشدار کاذب و دقت متفاوتی بوده‌اند، کارهای زیادی در زمینه افزایش دقت شناسایی حملات صورت گرفته است[۵][۱][۶][۷][۸][۹][۱۰] از آخرین کارهای انجام شده در این زمینه [۱][۱۱][۱۲][۸][۱۳] بوده‌اند که از روش های یادگیری عمیق استفاده کرده‌اند.

در مقاله [۱] برای بهبودی دقت تشخیص نفوذ از روش پروپسترون چندلایه (MLP) که یکی از روش‌های یادگیری عمیق است به‌کار برده است و در لایه اول و دوم از تابع فعالساز[۱] RELU استفاده می‌کنند که مشکل مرگ RELU دارد یعنی وقتی ورودی‌ها به صفر نزدیک می‌شوند یا صفر هستند گرادیان تابع صفر می‌شود بنابراین برای حل این مشکل در این پژوهش پیشنهاد شده است که از تابع فعالساز[۱۴] LEAKY RELU که با معرفی یک پارامتر α که اجازه می‌دهد گرادیان‌های کوچک در صورت عدم فعالسازی، فعال شوند و مزایای استفاده از این تابع از لحاظ محاسباتی سریع‌تر است و همینطور استفاده از الگوریتم بهینه[۱۴] ADAM به جای استفاده از الگوریتم گرادیان کاهشی[۱] که مهمترین نقطه ضعف این الگوریتم انتخاب درست نرخ یادگیری است که کار آسانی نیست را با الگوریتم[۱۴] ADAM که دارای نرخ یادگیری تطبیقی است شبکه را سریع‌تر همگرا کرد و با اضافه کردن یک لایه پنهان به شبکه، تعداد لایه‌ها از ۳ لایه به ۴ لایه تغییر پیدا می‌کند که معماری شبکه شامل یک لایه ورودی، دو لایه پنهان و یک لایه خروجی می‌باشدکه با عمیق شدن شبکه باعث کارآمدتر شدن شبکه می‌شود.

باتوجه به ماهیت الگوریتم بهینه‌ساز[۱۴] ADAM که سرعت همگرایی زیادی دارد و مشکل کاهش نرخ یادگیری را حل می‌کند و ماهیت تابع فعالساز[۱۴] LEAKY RELU که مشکل مرگ نرون‌ها را از بین می‌برد و افزایش تعداد لایه‌های مخفی شبکه، پیش‌بینی می‌کنیم که مدل پیشنهادی دقت بیشتری نسبت به آخرین کار انجام شده [۱] داشته باشد.

۱-۳ اهمیت و هدف و جنبه های نوآوری پژوهش

به طور کلی اهمیت اصلی پژوهش به منظور بهبود دقت تشخیص نفوذ در شبکه‌های کامپیوتری می‌باشد که نوآوری این پژوهش برای حل این مسأله استفاده از یادگیری عمیق می‌باشد با توجه به اینکه حملات و روش‌های نفوذ از سوی هکرها هرروزه در حال پیشرفت و گسترش می‌باشد. فعالیت در این حوزه نیز باتوجه به اهمیت بالا اطلاعات در حوزه‌های نظیر نظامی، پزشکی، سیاسی و بانکداری و امور مالی نیازمند شناسایی و روش‌های مقابله با آن بسیار هائز اهمیت است. برای بهبود

چالش دقت در این شبکه‌ها طراحی شبکه پروسپترون را با چهار لایه که شامل یک لایه ورودی و دو لایه پنهان و یک لایه خروجی انجام شده است تا یادگیری را کارآمدتر کند و از الگوریتم ADAM[۱۴] برای به روزرسانی وزن‌ها و رسیدن به تابع هدف و استفاده از تابع فعال ساز LEAKIY RELU[۱۴] که مشکل مرگ نرون ها را رفع می‌کند استفاده کرده است تا به هدف پرچالش این حوزه بهبود دقت سیستم تشخیص نفوذ کمکی شود. در نهایت با افزایش دقت به۹۳ درصد کارایی مدل پیشنهادی در مقابل سایر روش‌ها به اثبات می‌رسد.

۱-٤ فرضیه های پژوهش

پیش بینی می‌کنیم که با افزایش دادن لایه‌های مخفی از یک لایه به دو لایه و استفاده از الگوریتم ADAM [۱۴] جهت یادگیری خودکار نرخ یادگیری که سبب افزایش سرعت همگرایی می‌شود و استفاده از تابع LEAKY RELU [۱۴] در نرون‌های لایه ورودی و مخفی که از مرگ نرون‌ها جلوگیری می‌کند در طراحی شبکه‌های پروسپترون چند لایه و شاهد بهبود معیار دقت در این شبکه‌ها باشیم.

۱-۵ دامنه تحقیق

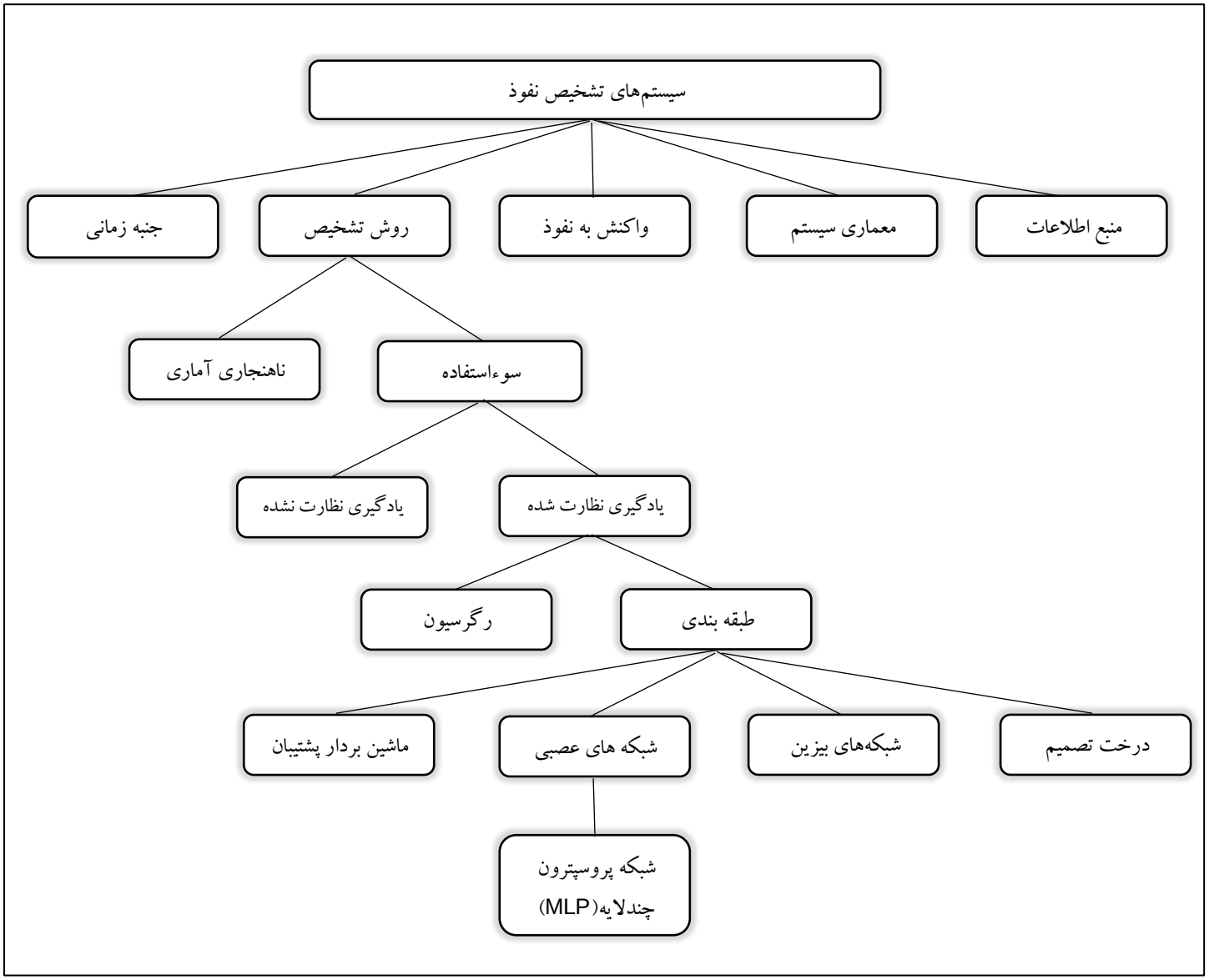
دامنه تحقیق در شکل ۱-۲ نمایش داده شده‌است. همانطور که در شکل ۱-۲ دیده می‌شود تحقیقات در سیستم‌های تشخیص نفوذ از پنج دیدگاه یک مبتنی بر منبع اطلاعات(Information Source)، دوم معماری سیستم (Architecture)، سوم واکنش به نفوذ (Architecture)، چهارم روش تشخیص (Detection Methodology) و پنجم جنبه زمانی (Time Aspect) قابل بررسی می‌باشد.

از دیدگاه توع روش تشخیص شامل دو تکنیک است که یکی مبتنی بر تطابق الگو یا سوءاستفاده در این روش از الگوهای حملات شناخته‌شده، برای شناسایی و تشخیص نفوذ استفاده می‌کند. در این روش الگوهای نفوذ از پیش ایجاد شده و به‌صورت قانون نگهداری می‌شوند و روش دوم ناهنجاری آماری که این روش رفتار کاربر را با رفتار نرمالی که در پایگاه داده ذخیره شده است، مقایسه می‌کند و از آن برای تشخیص حملات شناخته‌شده استفاده می‌کند و با استفاده از روش آماری سعی در پیدا کردن فعالیت‌هایی دارد که با الگوی رفتار نرمال تطابق ندارند و ناهنجار و غیر نرمال به نظر می‌رسند.

برای تشخیص نفوذ معمولاً از الگوریتم‌های داده‌کاوی جهت تشخیص حمله استفاده می‌شود. الگوریتم‌های مربوطه نیز به دو گروه نظارت‌شده و نظارت‌نشده طبقه‌بندی می‌شوند. در سیستم‌های تشخیص نفوذ معمولاً از روش یادگیری ماشین به شکل طبقه‌بندی یا خوشه‌بندی استفاده می‌شود، که پیش‌بینی برای داده‌های آینده بر اساس داده‌های گذشته است این روش در مجموعه یادگیری نظارت‌شده جای دارد. میتوان از الگوریتم‌های معروف در این خصوص ماشین بردار پشتیبان(SVM[۱]) درخت تصمیم(DT[۲])، شبکه‌های بیزین(BN[۳])، شبکه‌های عصبی مصنوعی(NN[۴]) را نام برد[۱۵]از جمله پرکاربردترین

[1] Support vector machine
[2] Decision Trees
[3] Baysian Networks
[4] Naural Networks

الگوریتم‌های شبکه‌های عصبی می‌توان به الگوریتم شبکه‌های پروسپترون چند لایه(MLP[1])، شبکه‌های عصبی بازگشتی(RNN[2]) ، شبکه‌های خود رمزگذار(AE[3]) و غیره می‌توان نام برد.در این تحقیق ما در زیر شاخه شبکه‌های پروسپترون چندلایه تلاش می‌کنیم تا با الگوریتم[۱۴] ADAM بتوانیم دقت را در سیستم‌های تشخیص نفوذ [۱] بهبود دهیم.



شکل ۱- ۲. دامنه تحقیق

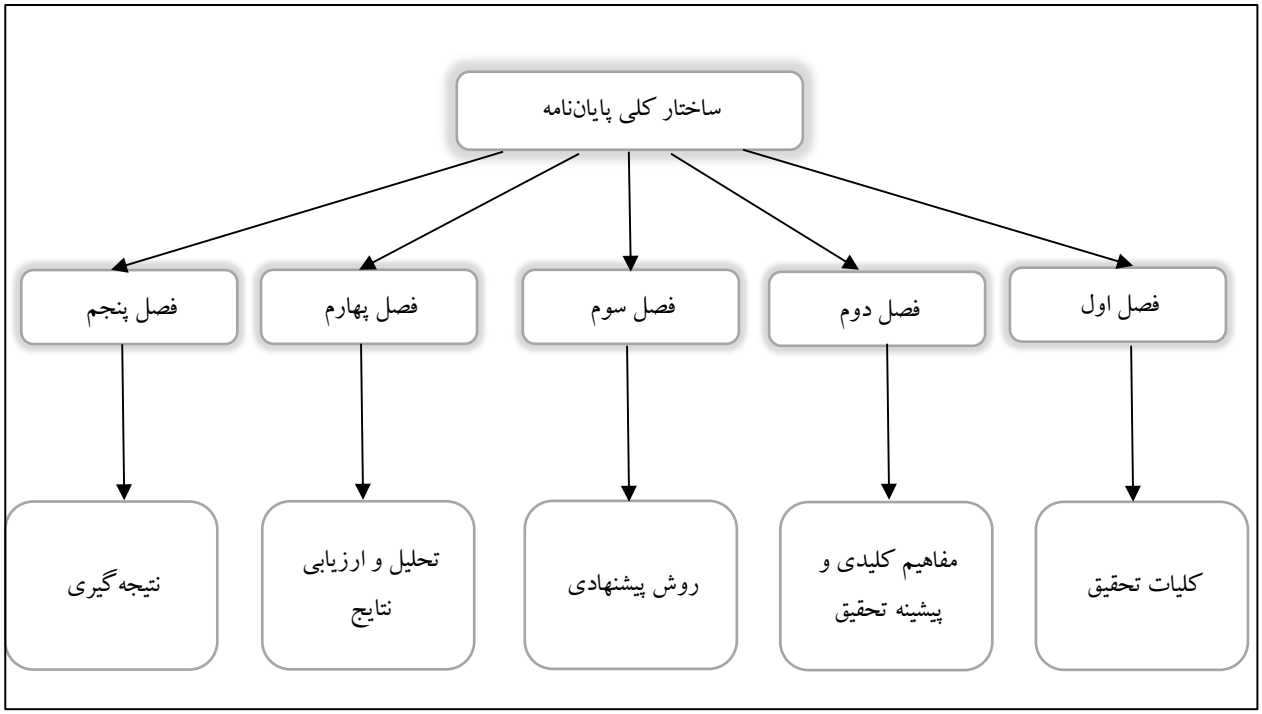
[1] Multi Layer Perceptron
[2] Recurrent Neural Network
[3] Auto Encoder

۱-۵ مختصری بر روش تحقیق

در ابتدا مجموعه داده NSL-KDD را با روش تحلیل مولفه‌های اساسی(PCA) که یکی از روش‌های کاهش ویژگی به روش استخراج ویژگی می‌باشد، کاهش ابعاد می‌دهیم و در مرحله بعدی مجموعه داده را به شبکه پروسپترون ۴ لایه که با الگوریتم ADAM[۱۴] جهت به روزسانی وزن‌ها و استفاده از تابع فعالساز LEAKY RELU[۱۴] در لایه اول، دوم و سوم طراحی کرده‌ایم تزریق می کنیم و در نهایت با مقایسه مقدار بدست آمده با تابع هدف از طریق تابع ضرر دقت را محاسبه می کنیم و نسبت به کار پایه [۱]ارزیابی می کنیم

۱-۶ ساختار پایان نامه

درشکل ۱-۳ یک ساختار کلی از پایان‌نامه را نمایش داده‌ایم در بخش اول از فصل دوم سعی در آشنایی خوانندگان با سیستم‌های تشخیص نفوذ و مقدمه ای بر یادگیری عمیق و مفاهیم کلیدی این شبکه با یک معرفی کلی و اجمالی پرداخته شده است و در بخش دوم از این فصل به کارهای انجام شده در این زمینه می‌پردازیم و در فصل سوم به ارائه استراتژدی پیشنهادی پرداخته خواهد شد و نوآوری پژوهش در این بخش مطرح می‌شود و در فصل چهارم از این پایان‌نامه به انجام آزمایش در راستای این استراتژی پیشنهادی و بررسی نتایج حاصل از آن مورد بررسی قرار خواهد گرفت و در فصل پایانی جمع‌بندی نهایی ذکر خواهد شد.



شکل ۱- ۳. ساختار کلی پایان نامه

۲.پیشینه تحقیق و مروری بر کارهای انجام شده

۲-۱ مقدمه

با افزایش روز افزون شبکه‌های کامپیوتری و ایجاد مشکلات امنیتی، روش‌های جهت تشخیص، جلوگیری و مقابله با حملات ارائه شده است که در بحث تشخیص با عنوان سیستم‌های تشخیص نفوذ جهت شناسایی تهدیدات عنوان می‌شود.که یا توجه به تغییرکردن نوع حملات در شبکه‌ها و پایگاه داده‌ها این حوزه نیازمند به‌روزرسانی هرروزه در جهت بهبود عملکرد خود می‌باشد. یکی از مسائل مهم در این حوزه بهبود دقت تشخیص این سیستم‌ها می‌باشد از این‌رو تحقیقات زیادی در صورت گرفته است. در ادامه این فصل در ابتدا به اصطلاحات رایج در زمینه شبکه‌های عصبی می‌باشد می‌پردازیم و در بخش بعدی به کارهای انجام شده توسط محققان در این زمینه که به اختصار توضیح داده خواهد شد. و در انتها یک جمع‌بندی از این فصل را خواهیم داشت.

۲-۲مروری برسیستم تشخیص نفوذ

از یک سو افزایش روزافزون حملات به شبکه‌های کامپیوتری و از سوی دیگر وابستگی شدید فعالیت‌های حوزه‌های مختلف صنعتی و بازرگانی و نظامی و پزشکی و آموزشی و غیره به سرویس‌های که توسط شبکه‌های کامپیوتری ارائه می‌شود تلاش برای جلوگیری و نفود به این شبکه‌ها امری حیاتی و ضروری می‌باشد.

تشخیص نفوذ به این معنا می‌باشد که فرایند شناسایی و پاسخ به فعالیت‌های مخرب که به صورت هدفمند حمله می‌کنند وجود سیستم تشخیص نفوذ کارآمد بسیار در بحث امنیت مهم تلقی می‌شود.راه‌های نفود به هر شبکه تا حد زیادی به تشخیص و شناسایی نقاط ضعف آن شبکه بستگی دارد.نوع طراحی و اشکالات نرم‌افزاری داخل شبکه موجب نفود به آن شود و امنیت شبکه را تهدید کند.

در سیستم‌های تشخیص نفوذ، تشخیص رفتارهای غیرعادی با جع‌آوری مقدار کافی از داده‌ها، تجزیه و تحلیل شبکه‌ها قابل بررسی می‌باشد.سیستم‌های تشخیص نفوذ سبب شناسایی و مستند ساری تهدیدات موجود شده و مانع از شکل‌گیری کامل یک حمله می‌شود و می‌تواند با هشدار از حملات مشابه جلوگیری کند.

به‌طور کلی هدف سیستم‌های تشخیص نفوذ را می‌توان به صورت زیر خلاصه کرد:

- امکان برقراری ارتباط بین یک واقعه و شخص مسئول آن واقعه
- امکان شناسایه حمله و واکنش قابل قبول برا مقابله یا توقف آن و جلوگیری از تکرار حمه
- فرایند تولید، ثبت و مرور یک سابقه از عملکرد سیستم

سه عملکرد پایه می‌توان برای سیستم‌های تشخیص نفود در نظر گرفت:

۱. نظارت و ارزیابی رفتار شبکه
۲. کشف و تشخیص الگوهای منطبق با حمللات شناخته شده
۳. واکنش و تحلیل الگوی رفتارهای ناهنجار

بر همین اساس هر سیستم تشخیص نفوذ را می توان بر اساس روش‌های تشخیص نفوذ، معماری سیستم، منبع اطلاعات، واکنش به نفوذ و جنبه زمانی دسته‌بندی کرد.

۲-۲-۱ تشخیص نفوذ

نفوذ فعالیتی است که توسط آن محرمانگی، صحت یا دسترسی پذیری به منابع دچار اختلال می‌شود. تشخیص نفوذ درواقع شناسایی دستیابی‌های غیر مجاز به حملات انجام شده به شبکه است. سامانه تشخیص نفوذ وظیفه نظارت بر فعالیت سامانه، تجزیه و تحلیل بسته‌های شبکه، تعیین الگوی حملات و ارزیابی صحت و یک‌پارچگی فایل‌ها را بر عهده دارد. معمولاً سامانه‌های تشخیص نفوذ با یک ساختار سه سطحی توصیف می‌شوند که عبارتند از:

- واحد نظارت:که وظیفه ثبت رخدادها را بر عهده دارد. رخدادهای سامانه توسط این واحد جهت تشخیص و تحلیل ثبت می‌شوند. همچنین نمایش هشدار به مدیر شبکه یا مسئول نظارت و پیگیری رویداد های شبکه بر عهده این واحد است.
- واحد تشخیص و تحلیل: اطلاعات ثبت شده توسط واحد نظارت را به عنوان ورودی دریافت می کند و سپس بر اساس این اطلاعات به یکی از روشهای مبتنی بر سوءاستفاده یا مبتنی بر ناهنجاری، مدل ساخته می‌شود. حملات و تهدیدات امنیتی به کمک مدل ساخته شده تمایز داده می‌شوند.
- واحد هشدار: وظیفه دارد که متناسب با نوع حمله رفتار مناسب را انجام دهد. این واحد در صورت نیاز اخطارهای مناسب را برای ثبت یا نمایش به واحد نظارت ارسال میکند[۱۶]، [۱۷]

۲-۲-۲ معماری سیستم‌های تشخیص نفوذ

سیستم های تشخیص نفوذ براساس مکان قرارگیری در سیستم شبکه و دامنه فعالیت به دو گروه سیستم های تشخیص نفوذ مبتنی بر میزبان (HIDS[1]) سیستم‌های تشخیص نفوذ مبتنی بر شبکه (NIDS[2]) طبقه‌بندی می‌شوند. معمولا برای دست‌یابی به حداکثر کارایی و امنیت این سیستم‌ها به صورت ترکیبی نیز مورد استفاده قرار می گیرند که با نام سیستم‌های تشخیص نفوذ توزیع شده (DIDS[3]) شناخته می‌شوند[۱۸]

۲-۲-۲-۱ سیستم های تشخیص نفوذ مبتنی بر میزبان(HIDS)

این سیستم‌ها شناسایی و تشخیص فعالیت های غیر مجاز بر روی کامپیوتر میزبان را به عهده دارند. سیستم‌های تشخیص نفوذ مبتنی بر میزبان می‌توانند حملات و تهدیداتی مانند دسترسی به فایل‌ها ، اسب‌های تراوا و... را روی کامپیوتر میزبان تشخیص دهند. سیستم‌های تشخیص نفوذ مبتنی بر میزبان تنها بر روی کامپیوتر میزبان و یا کامپیوترهای منفرد اجرا می‌شوند و از کل شبکه اطلاع ندارند. این نوع سیستم‌ها فقط بسته‌های ورودی و خروجی به یک کامپیوتر را بررسی و نظارت کرده و هنگام تشخیص نفوذ و یا فعالیت مشکوک به مدیر شبکه و یا کاربر کامپیوتر هشدار می دهد.[۱۸]

[1] Host intrusion detection systems
[2] Network intrusion detection systems
[3] Distributed intrusion detection systems

۲-۲-۲-۲ سیستم‌های تشخیص نفوذ مبتنی برشبکه(NIDS)

سیستم‌های NIDS به‌منظور شناسایی تهدیدات، به نظارت و تجزیه و تحلیل ترافیک درسراسر شبکه می‌پردازند. این سیستم‌ها فعالیت‌های مخربی مانند حملات انکار سرویس DOS اسکن پورت و غیره را در کل شبکه شناسایی می کنند. سیستم‌های NIDS به‌صورت سخت‌افزار یا نرم‌افزار برای پایش ترافیک عبوری و کامپیوترهای موجود در شبکه، در مکان یا مکان‌های خاصی از شبکه قرار داده می‌شوند، تا ترافیک شبکه را مورد تحلیل قرار داده و هنگامی که حمله یا رفتار غیرعادی تشخیص داده می‌شود یک پیغام هشدار برای مدیر شبکه ارسال نمایند.[۱۸]

۲-۲-۲-۳ سیستم‌های تشخیص نفوذ توزیع شده(DIDS)

سیستم‌های تشخیص نفوذ ترکیبی که با نام توزیع شده نیز شناخته می‌شوند از چندین سیستم تشخیص نفوذ مبتنی بر شبکه یا سیستم تشخیص نفوذ مبتنی بر میزبان یا ترکیبی از این دو نوع تشکیل شده‌اند. این سیستم‌ها از انعطاف‌پذیری بالایی برخوردار هستند و سطح امنیت را افزایش می‌دهند.[۱۸]

۲-۲-۳ روش‌های تشخیص نفوذ

به‌طورکلی روش‌های تشخیص نفوذ در میزبان و شبکه به دو دسته مبتنی بر سوءاستفاده یا امضاء و روش مبتنی بر ناهنجاری طبقه‌بندی می‌شود.در هر دو روش تشخیص نفوذ مبتنی بر امضاء و تشخیص نفوذ مبتنی بر ناهنجاری از روش‌ها و الگوریتم‌های داده‌کاوی به‌منظور تشخیص حملات استفاده می‌شود. الگوریتم‌های داده‌کاوی و یادگیری ماشین به دو گروه عمده نظارت‌شده و نظارت‌نشده طبقه‌بندی می‌شوند. درزمینه‌ی تشخیص نفوذ، یادگیری ماشین عبارت است از یک روش یادگیری کامپیوتری مبتنی بر طبقه‌بندی یا خوشه‌بندی که بر اساس داده‌های گذشته برای داده‌های آینده که به سیستم وارد می‌شود، پیش‌بینی انجام می‌دهد. در یادگیری نظارت‌شده از ابتدا دسته‌ها مشخص هستند و هر یک از داده‌های آموزشی به دسته‌های خاص نسبت داده‌شده است. در این نوع یادگیری مجموعه داده آموزشی در دسترس است که فرآیند یادگیری بر اساس آن انجام می‌شود. روش طبقه‌بندی ازجمله روش‌های یادگیری نظارت‌شده است. برای تشخیص نفوذ در شبکه می‌توان از الگوریتم‌های طبقه‌بندی مانند شبکه‌های عصبی، نزدیکترین همسایه، بیزین، درخت‌های تصمیم و غیره جهت آموزش سیستم برای تشخیص حمله‌های از قبل شناخته‌شده استفاده کرد[۱۸]

۲-۲-۳-۱ روش تشخیص نفوذ مبتنی بر سوءاستفاده

این روش از الگوهای حملات شناخته‌شده، برای شناسایی و تشخیص نفوذ استفاده می کند. در این روش الگوهای نفوذ از پیش‌ایجاد شده و به‌صورت قانون نگهداری می‌شوند از این‌رو برای تشخیص حملاتی که الگوهای حملات آن در پایگاه داده‌ی الگوهای حمله ذخیره‌شده، روش مناسبی است. اشکال این روش ناتوانی در تشخیص حملات جدیدی است که الگوی آن در پایگاه داده مبتنی بر امضاء وجود ندارد و برای تشخیص این نوع حملات باید پایگاه داده با امضای حملات جدید به‌روزرسانی شود. ازآنجاییکه درروش مبتنی بر امضاء نیاز به یادگیری از محیط وجود ندارد، پیاده‌سازی این روش ساده است.

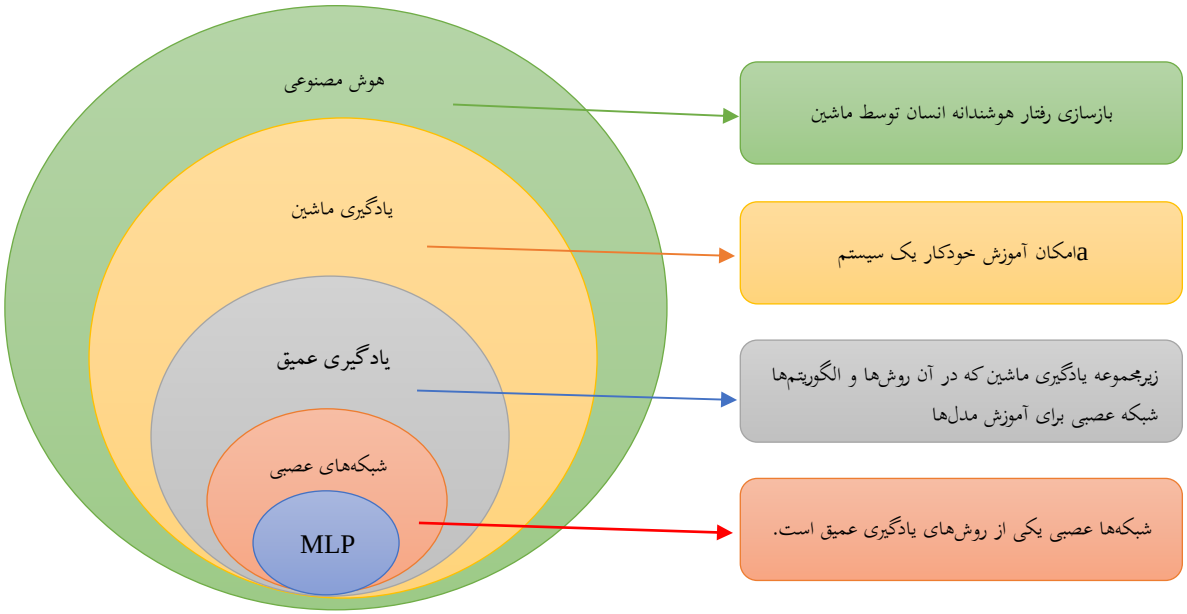
این روش بر اساس این معماری از پیش‌پردازش برای پیدا کردن و مقایسه الگوی فعالیت مشاهده‌شده در محیط شبکه با امضاءهای موجود در پایگاه داده سیستم تشخیص نفوذ استفاده می کند. اگر بین فعالیت مشاهده‌شده در شبکه با امضاءهای موجود در پایگاه داده تطبیقی مشاهده شود یک پیام هشدار صادر میشود. به دلیل اینکه روش تشخیص نفوذ مبتنی بر امضا حملات شناخته‌شده را باقابلیت اطمینان بالا و نرخ هشدار نادرست پایین شناسایی می کند. در بسیاری از سیستم‌های تشخیص نفوذ تجاری از این روش استفاده می‌شود[۱۸]

۲-۲-۳-۲ روش تشخیص نفوذ مبتنی بر ناهنجاری

این روش رفتار کاربر را با رفتار نرمالی که در پایگاه داده ذخیره‌شده است، مقایسه می کند و از آن برای تشخیص حملات شناخته‌شده استفاده می کند و با استفاده از روش آماری سعی در پیدا کردن فعالیت‌هایی دارد که با الگوی رفتار نرمال تطابق ندارند و ناهنجار و غیر نرمال به نظر می‌رسند. درواقع برای تشخیص ناهنجاری باید الگوهای خاصی را پیدا کرد و رفتارهایی که از الگو پیروی می کنند، نرمال و رویدادهای انحرافی به‌عنوان ناهنجاری تشخیص داده می‌شوند .در این روش مسئله مهم ایجاد نما از رفتار نرمال است زیرا رفتار نرمال کاربران ممکن است تغییر کند و لذا سیستم تشخیص نفوذی که از این روش استفاده می کند باید خود را با این تغییرات بروز کند. مسئله دیگری که در ساخت مدل رفتار نرمال وجود دارد انتخاب ویژگی‌هایی است که به‌عنوان ورودی برای ساخت مدل از آن استفاده می‌شود. در مدل‌های فعلی پارامترهای ورودی توسط کارشناس امنیت تعیین می‌شود و تضمینی وجود ندارد که همه ویژگی‌هایی که در تشخیص نفوذ مؤثر هستند به‌درستی انتخاب شوند. چنانچه ویژگی‌های مهم وابسته به نفوذ به اشتباه کنار گذاشته شوند تشخیص حمله از رفتار نرمال بسیار مشکل خواهد بود. همچنین عدم حذف ویژگی‌هایی که مرتبط با نفوذ نیستند می‌تواند باعث کاهش کارایی در تشخیص نفوذ شود. نقطه قوت این روش توانایی تشخیص حملات جدید است و ضعف این روش نرخ هشدار نادرست بالای آن است[۱۸]

۲-۳ مروری بر روش‌های و اصطلاحات رایج شبکه‌های عصبی جهت تشخیص نفوذ

اگرچه اصطلاح هوش مصنوعی، یادگیری ماشین و یادگیری عمیق در بیشتر اوقات به جای هم مورد استفاده قرار می گیرند و به یکدیگر مرتبط هستند، اما به‌طور کامل به موارد مشابهی اشاره نمی کنند. در شکل۲-۱ نحوی ارتباط آنها با یکدیگر قابل نمایش است و همانطور که مشاهده می‌شود، یادگیری عمیق زیرمجموعه‌های از یادگیری ماشین و همچنین هوش مصنوعی است.[۱۹]



شکل ۲- ۱.ارتباط بین هوش مصنوعی و یادگیری ماشین و یادگیری عمیق و شبکه‌های عصبی و شبکه‌های پروسپترون چندلایهmlp

در ادامه مروری بر هوش مصنوعی، یادگیری ماشین، روش‌های یادگیری عمیق و در ادامه اصطلاحات رایج که باید برای فعالیت در این زمینه با آن‌ها آشنا باشیم را به صورت مختصر توضیح خواهیم داد.

۲-۳-۱ هوش مصنوعی

همانطور که از نام آن پیداست هوش مصنوعی ترکیب هوش انسانی در ماشین است، به‌نحوی که رفتاری همانند انسان را تقلید و خلاقانه مسائل را حل کند. به‌طور دقیق‌تر هوش مصنوعی، در تلاش برای روشی است که رونوشتی از مغز انسان را پیاده‌سازی کند، یعنی همانگونه که یک انسان فکر می‌کند و عمل می‌کند.[۱۴]

۲-۳-۲ یادگیری ماشین

یادگیری ماشین که زیرمجموعه‌های از هوش مصنوعی است، رایانه را به‌گونه‌ای توانمند می‌سازد کع قادر به یادگیری از طریق تجربه و بدون برنامه‌ریزی صریح می‌شود. این یادگیری توسط داده‌های که مناسب به هر مسئله به آن داده می‌شود. رابطه‌های را میان ورودی و خروجی مسئله پیدا می‌کند تا در مواجه با مسئله‌های مشابه از آن استفاده کند. یادگیری ماشین اینگونه آموزش می‌بیند که، چگونه یک تصمیم‌گیری برای مسئله انجام دهد، و روشی برای تحقق بخشیدن به هوش مصنوعی است.[۱۴]

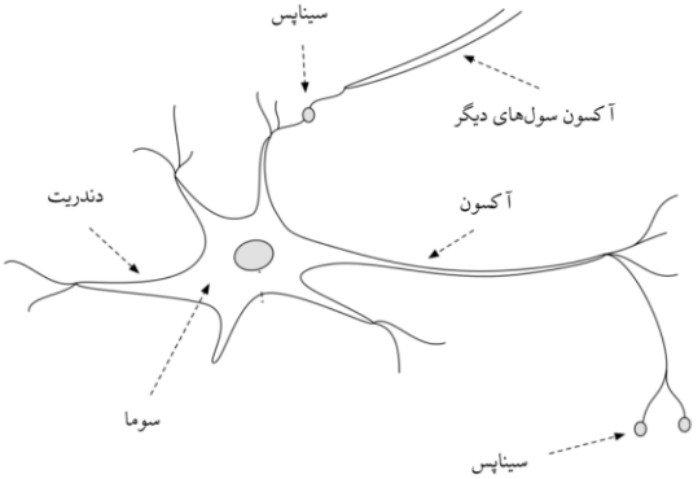
۲-۳-۳ یادگیری عمیق

یادگیری عمیق زیرمجموعه‌ای از یادگیری ماشین بوده و از ساختار شبکه‌های عصبی برای تقلید در تصمیم‌گیری حل یک مسئله مشابه مغز انسان استفاده می‌کند، و همان‌کار یادگیری ماشین را انجام می‌دهد، ولی قابلیت‌های متفاوتی در آن

وجود دارد. در مقایسه یادگیری ماشین با یادگیری عمیق اینگونه می‌توان بیان کرد، در حالیکه یادگیری عمیق به طور خودکار ویژگی‌ها را از ساختار داده‌ها استخراج می‌کند، این عملتوسط یادگیری ماشین باید به صورت دستی انجام گیرد[۱۴].

۲-۳-٤ شبکه‌های مصنوعی

شبکه‌های عصبی مصنوعی مدل‌های محاسباتی بوده که سازوکار یادگیری را همانند شبکه عصبی طبیعی از ساختار مغز انسان شبیه‌سازی می‌کند. ساختار مغز انسان که همان شبکه عصبی طبیعی است از تعداد زیادی واحد به نام نرون تشکیل شده است ساختار نرون در شکل۲-۲ نمایش داده شده است.



شکل ۲- ۲. نمایی از یک شبکه عصبی واقعی [١٤]

۲-۳-٤-۱ مقدمات شبکه‌های عصبی

در این بخش به معرفی مختصری از اصطلاحاتی که در این شبکه مورد نیاز است می‌پردازیم

۲-۳-٤-۲ نرون

نرون ها همان گره ها در طراحی شبکه می‌باشد که ورودی را می‌گیرند و بعد از پردازش به یک خروجی تبدیل می‌کند. در طراحی مدل شبکه عصبی تعداد نرون‌ها در هر لایه بسیار مهم است تعداد نرون‌های ورودی معمولا همان ویژگی ها و ابعاد مسئله می‌باشد و تعداد نرون‌ها در لایه خروجی همان تعداد خروجی مسئله می‌باشد و اما تعداد نرون‌ها در لایه مخفی متفاوت‌تر است در زیر سه روش جهت انتخاب صحیح تعداد نرون‌ها ارائه شده است باید به این نکته توجه داشت که انتخاب صحیح جهت کارایی مدل توسط تست تعداد آن به صورت عملی با آزمون و خطا در سیستم بدست می‌آید اما خب در ابتدا اینکه تصادفی این تعداد انتخاب شود بسیار زمان بر و سخت می‌باشد

روش اول با در نظر گرفتن S نرون در لایه مخفی، می‌توان 2^S کلاس را طبقه‌بندی کرد.

روش دوم که نسبتا کاربردی نیز است به صورت زیر می باشد

$$N_n = \sqrt{(N_i + N_o)}$$

(۲-۱)

N_n: تعداد مناسب نرون در لایه پنهان

N_i: نرون‌های لایه ورودی

N_o : نرون‌های لایه خروجی

روش سوم: روش دیگری جهت تخمین تعداد نرون ها در لایه پنهان در معادله (۲-۲) نمایش داده شده است که a عامل مقیاس پذیری دلخواه بین(۱۰-۲)می‌باشد.

$$N_n = \frac{N_s}{\left(a * (N_i + N_o)\right)}$$

(۲-۲)

N_n: تعداد نرون‌ها در هر لایه

N_s: تعداد نمونه‌های موجود در در داده‌های آموزشی

N_o: تعداد نرون در خروجی

N_i: تعداد نرون‌های ورودی

α: عامل مقیاس‌بندی

۲-۳-٤-۳ وزن

وزن یک متغییر می‌باشد زمانی‌که مقادیر ورودی به نرون اعمال می‌شود از طریق ضرب در نرون، آن نرون، وزن‌دار می‌شود. ایده اصلی شبکه‌های عصبی این است که با تغییر مقادیر وزن و بایاس شبکه یک رفتار بهینه در جهت تابع هدف به ما بدهد. مقداردهی اولیه وزن‌ها برای دستیابی به همگرایی شبکه بسیار مهم است که روش‌های مختلفی از جمله مقداردهی اولیه ثابت، مقداردهی اولیه نرمال یا یکنواخت، مقداردهی LeCun نرمال و یکنواخت، مقداردهی Glorot یکنواخت و نرمال، مقداردهی He می‌توان اشاره کرد.

۲-۳-٤-٤ بایاس

بایاس یکی دیگر از اجزاء خطی که روی ورودی تاثیر می‌گذارد که به حاصل‌ضرب ورودی در وزن اضافه می‌شود.یک متغییر قابل تنظیم برای نرون ها می‌باشد و در هر لایه ممکن است مقادیر متفاوتی داشته باشد.

۲-۳-٤-۵ توابع فعالساز

یکی از اجزا بسیار مهم درمعماری یک مدل از شبکه‌های عصبی توابع فعاساز می باشد که در واقع راهی برای معرفی غیرخطی می‌باشد. از این توابع جهت تعیین خصوصیات نرون در راستای حل مسائل استفاده می‌شود که با توجه ماهیت هرتابع برای مسئله انتخاب می‌شود و وظیفه انتشار خروجی از یک لایه به لایه دیگر بعد از انجام مرحله محاسباتی در هر نرون را به‌عهده دارد.از توابع پرکاربرد می‌توان به Relu، Sigmoid، Than، Softmax، و... می‌توان نام برد.

۲-۳-٤-٦ لایه‌های ورودی و خروجی و مخفی

در طراحی شبکه‌های عصبی یک لایه جهت دریافت ورودی‌ها لایه اول شبکه می‌باشد و یک یا چند لایه مخفی جهت پردازش با توجه به پیچیدگی مساله انتخاب می‌شود باید توجه داشت که افزایش بی‌مورد لایه مخفی بار محاسباتی را افزایش می‌دهد و اینکه این لایه برای ما قابل مشاهده نمی‌باشد و در انتها انتخاب یک لایه خروجی که خروجی را تولید می‌کند برای ما قابل مشاهده می‌باشد در نظر گرفته می‌شود.

۲-۳-٤-۷ مدل‌های شبکه عصبی

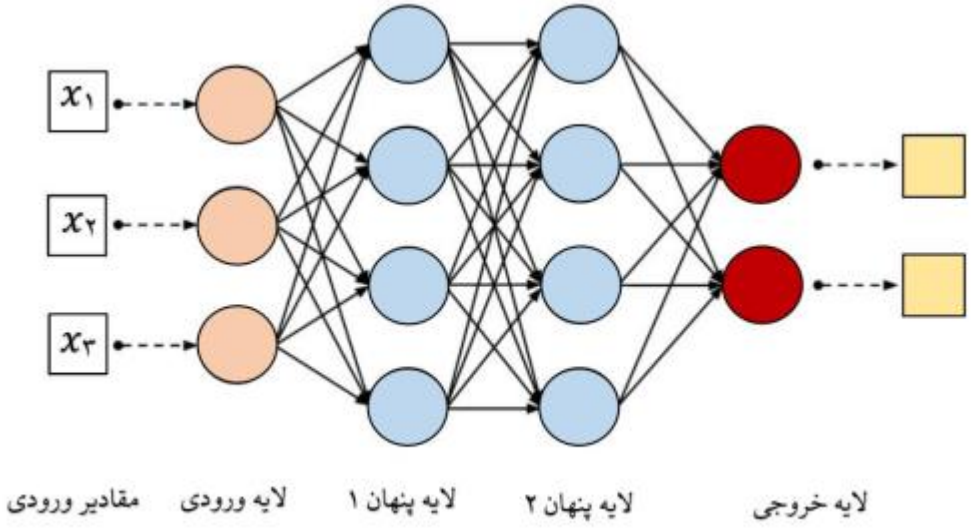
بهترین تعریف از شبکه‌های عصبی را Liping Yang ارائه داده است به این صورت که« شبکه‌های عصبی از چندین نورون مصنوعی تشکیل شده است که اطلاعات را بین یکدیگر رد و بدل می‌کنند، و هر کدام وزن‌هایی دارند که براساس «تجربه‌ی» شبکه شکل می‌گیرند. نورون‌ها یک نقطه‌ی فعال‌سازی دارند که اگر مجموع وزن و داده‌های ارسال شده به آن‌ها از آن نقطه عبور کند، آن‌ها فعال می‌شوند. مجموعه‌ای از نورون‌های فعال شده منجر به «یادگیری» می‌شوند.»

شبکه‌های عصبی انواع مختلفی از جمله شبکه‌های پروسپترون چندلایه(MLP)، شبکه‌های بازگشتی(RNN) و غیره دارد.

۲-۳-٤-۸ شبکه‌های پروسپترون چندلایه

شبکه‌های پروسپترون چند لایه که باعنوان شبکه‌های پیش‌خور عمیق نیز یاد می‌شود دو یا چند نرون می‌توانند باهم در قالب یک لایه ترکیب شوند و یک شبکه خود می‌تواند از چند لایه تشکل شده باشد در شبکه های پروسپترون هر نرون در هر لایه به تمام نرون‌ها لایه بعدی متصل است و به اصطلاح اتصال در این شبکه‌ها به صورت تمام متصل است.

در این شبکه‌ها با رفتن به هر لایه دیگر، جمع وزن‌دار مجموعه نرون‌های لایه قبلی محاسبه شده و پس از اعمال تابع فعالساز غیرخطی به لایه دیگر منتقل می‌شوند تا در نهایت به لایه خروجی برسند. در شکل ۲-۳ شمایی از یک شبکه عصبی پیشخور عمیق تماما متصل قابل مشاهده است.[۱۴]



شکل ۲- ۳. معماری یک شبکه پروسپترون تمام متصل [۱٤]

۲-۳-٤-۹ روش یادگیری

روش‌های یادگیری در شبکه‌های عصبی را می توان به سه نوع اصلی نظارتی، غیرنظارتی، تقویتی طبقه‌بندی کرد که در یادگیری نظارتی برای هر الگوی ورودی که در آموزش شبکه استفاده می‌شود یک الگوی خروجی مطلوب یا همان هدف در نظرگرفته می‌شود. یادگیری غیرنظارتی خروجی هدف در شبکه مشخص نمی‌باشد و روش تقویتی خروجی مطلوب مشخص نمی‌باشد بلکه فقط مشخص می کند خروجی درست است یا نه و در پاسخ‌های درست پاداش در مقابل پاسخ‌های اشتباه جریمه می‌شوند. نمونه یادگیری می توان به Hebbian، یادگیری کاهش گرادیان، یادگیری رقابتی، یادگیری تصادفی اشاره کرد.

۲-۳-٤-۹-۱ یادگیری کاهشی گرادیان

این روش بر اساس کمینه کردن خطایE ، که بر حسب وزن‌ها فعالسازی شبکه بیان می‌شود، عمل می کند. همچنین لازم است تابع فعالسازی که در شبکه استفاده می‌شود ، قابل مشتق‌گیری باشد. چون به‌روزرسانی وزن‌ها بر اساس گرادیان خطای E است بنابراین ، اگر Δw_{ij} وزن به‌روز شده لینک نرون jامین دولایه مجاور باشد. Δw_{ij} به صورت زیر تعریف می‌شود

$$\Delta w_{ij} = \eta \frac{\partial E}{\partial W_{ij}}$$

(۳-۲)

که در آن η پارامتر نرخ یادگیری و $\partial E / \partial W_{ij}$ گرادیان خطاست که نسبت به وزن W_{ij} محاسبه می‌شود. روش‌های دلتای Hoffو Windrowو یادگیری پس انتشار مثال‌هایی از این مکانیزم یادگیری هستند.[۱۹]

۲-۳-٤-۱۰ انتشار رو به جلو

به حرکت ورودی از لایه مخفی به سمت لایه خروجی یعنی ابتدا داده‌ها وارد لایه ورودی و سپس بعد از گذر از لایه‌های مخفی به لایه خروجی می‌رسند را انتشار رو به جلو گفته‌می‌شود

۲-۳-٤-۱۱ انتشار به عقب

در ابتدای طراحی شبکه مقدار وزن‌ها و بایاس را با توجه به مسئله یکی از روش‌های مقدار دهی اولیه در نظر می گیریم زمانی که خروجی مدل را می گیریم و با تابع هدف مقایسه و توسط تابع هزینهاین مقدار را پیش‌بنی می کنیم سپس این مقدار را به شبکه بر می گردانیم تا مقدار وزن‌ها و بایاس به‌روزرسانی شوند این وزن‌ها به گونه‌ای به‌روزرسانی می‌شوند که خطای شبکه را کاهش دهد.به این عمل انتشار رو به عقب گفته می‌شود.

۲-۳-٤-۱۲ تابع هزینه

یک شبکه را به‌گونه ایی طراحی می کنیم که خروجی دقیقی را پیش‌بینی کند میزان دقت شبکه توسط تابع هزینه بدست می‌آید.این تابع به نام‌های تابع ضرر یا تابع زیان هم شناخته شده استف هدف از طراحی یک شبکه این است که دقت پیش‌بینی را افزایش و خطاها را کاهش دهیم. نمایش مقدار کم این تابع نشان از عملکرد بهینه شبکه می‌دهد.

۲-۳-٤-۱۳معیار عملکرد

برای ارزیابی عملکرد الگوریتم ها ، معیار کمی برای باید تعریف شود نمونه ای از آن‌ها صحت، دقت، هشدار نرخ کاذب و بار محاسباتی و... است که با آن‌ها عملکرد شبکه را مورد ارزیابی قرار می‌دهیم.

۲-۳-٤-۱٤ بهینه‌سازی وزن‌ها

بهینه‌سازها الگوریتم‌هایی هستند که از طریق به‌هنگام سازی وزن‌ها در شبکه سعی دربه حداقل رساندن تابع زیان دارند، یعنی همان هدف اصلی ما از آموزش شبکه‌ها برای مسئله موردنظر که سعی داریم، وزن‌های شبکه به‌گونه‌ای تنظیم تا شبکه توانایی یادگیری را بدست آورد. انتخاب الگوریتم بهینه‌سازی مناسب نقش مهمی درسرعت همگرایی شبکه دارد. روش‌های مختلفی برای بهینه‌سازی وجود دارد. [۱۴]

گرادیان کاهش یکی از محبوب‌ترین و متداول‌ترین این الگوریتم‌ها در شبکه عصبی است نمونه‌های دیگر می‌توان به روش نیوتن، Adam و.... می‌توان نام برد.

۲-۳-٤-۱۵ نرخ یادگیری

میزان کاهش هزینه در هر تکرار را نرخ یادگیری می‌نامند یا به عبارتی دیگر سرعت کاهش هزینه که باید به دقت مشخص شود چرا که انخاب مقدار زیاد ممکن است حالت بهینه را رد کند و مقدار کم آن سبب طولانی شدن روند یادگیری شبکه شود.

۲-۳-٤-۱٦ بسته‌ها

در زمان آموزش شبکه به جای اینکه کل ارسال کل ورودی آن‌ها را به بسته‌های کوچک با اندازه یکسان تقسیم می‌کنیم این‌کار سبب جامع‌تر از مدلی که تمام داده‌ها بهصورت یکجا به شبکه تزریق شود دارد.

۲-۳-٤-۱۷ دوره‌ها

به انتشار روبه جلو و رو به عقب در شبکه یک دوره می‌گوییم که تنظیم تعداد آن توسط خود طراح صورت می‌گیرد و برای یادگیری شبکه به تکرارهای مناسب احیتاج است اما باید توجه داشت که تعداد زیاد دوره ها سبب overfit شدن شبکه می‌شود.

۲-۳-۵ چالش های آموزش در شبکه‌های عمیق

فرآیند آموزش در شبکه‌های عمیق شامل یافتن مجموعه‌ای از وزن‌ها در شبکه است، این وزن‌ها نمایانگر یادگیری در شبکه برای مسئله مورد نظر هستند آموزش در شبکه‌های عصبی روندی تکرار شونده است، یعنی گام‌به‌گام با بروزرسانی‌های کوچک در وزن‌ها و با تکرار این عمل، عملکرد مدل در حل مسئله بهبود پیدا می‌کند. این فرآیند، یک مسئله بهینه‌سازی را بوجود می‌آورد، چرا که شبکه سعی می‌کند براساس وزن‌ها تابع زیان را حداقل کند. همین مسئله بهینه‌سازی سبب ایجاد چالش‌هایی در شبکه می‌شود. سوال اینجاست که دقیقا چه چیزی در این مسئله بهینه‌سازی چالش برانگیر میباشد؟[۱۴]

۲-۳-۵-۱ محو گرادیان و انفجار گرادیان

هنگام استفاده از الگوریتم پس‌انتشار، در فاز عقب‌گرد محاسبه گرادیان کوچک و کوچک‌تر می‌شود. این اتفاق به این دلیل بوجود می‌آید که، گرادیان کاهشی در هر تکرار مشتقات جزئی را با طی کردن از لایه پایانی به سمت لایه ابتدایی با استفاده از قانون زنجیرهای می‌یابند. در شبکه‌ای با داشتن n لایه پنهان، مشتقات این n لایه در یکدیگر ضرب می‌شود. حال اگر این مشتقات کوچک باشند، با رفتن به لایه‌های اولیه به صورت نمایی کاهش پیدا و (یا در بدترین حالت صفر می‌شوند و یادگیری شبکه متوقف می‌شود) همین امر سبب پدیده محو گرادیان می‌شود. از آنجایی که این گرادیان کوچک در تکرار الگوریتم به‌هنگام‌سازی نمی‌شوند و این لایه‌های اولیه اغلب در شناخت داده‌ها موثر هستند، منجر به عدم دقت کافی شبکه می‌شوند. در مقابل این اگر این مقادیر مشتقات بزرگ باشند با رشد نمایی از طریق انتقال به لایه‌ها، سبب سرریز شده و وزن‌ها دیگر توانایی به‌هنگام‌سازی نخواهند داشت، و شبکه‌ای ناپایدار را پدید می‌آورد که با عنوان انفجار گرادیان شناخته می‌شود.[۱۴]

۲-۳-۵-۲ شناسایی محو گرادیان و انفجار گرادیان

راه‌های شناسایی هرکدام از این چالش‌ها در زیر شرح داده شده است:

شناسایی انفجار گرادیان

- به دلیل عدم ثبات مدل، تغییرات زیادی در به‌هنگام‌سازی وزن‌ها مشاهده شود. وزن‌ها در هنگام آموزش به‌صورت نمایی رشد می‌کنند.
- در طول فرآیند آموزش تابع هزینه مقدار NaN بگیرد.
- مدل اطلاعات زیادی را در طول فرآیند آموزش فرا نمی‌گیرد، بنابراین تابع هزینه ضعیفی دارد.

شناسایی محو گرادیان

- در طول فرآیند آموزش بهبود مدل بسیار آهسته می‌باشد، و ممکن است فرآیند آموزش خیلی زود متوقف شود یعنی، هیچ آموزش دیگری سبب بهبود مدل نمی‌شود.
- وزن‌های نزدیک به لایه خروجی شاهد تغییراتی بیشتری نسبت به لایه‌های نزدیک به ورودی دارند.
- وزن‌های مدل به‌صورت نمایی کاهش پیدا کند.
- وزن مدل در هنگام آموزش صفر شود.

با توجه به مشکلاتی که در حین آموزش شبکه روبه‌رو می‌شویم می‌توانیم شناسایی کنیم که درگیر انفجار گرادیان شدیم یا محوگرادیان که با توجه به آن راهی را جهت مقابله انتخاب کنیم.

۳-۲-۶ تفاوت پارامترها و ابرپارامترها

در آموزش شبکه‌های عصبی دو نوع پرامتر نقش دارند که با یکدیگر متفاوت هستند در زیر به صورت خلاصه ایی در مورشان توضیح خواهم داد

پارامترهای مدل: مقادیری هستند که از مجموعه داده‌ها در حین آموزش خود شبکه تخمین می‌زند و به صورت دستی قابل تنظیم نمی‌باشد و متغییرهای داخلی شبکه هستند. مدل از این پرامترها جهت پیش‌بینی استفاده می‌کند.
ابرپارامترها: متغییرهای خارجی در در مدل‌سازی شبکه هستند که قبل از شروع یادگیری تعیین می‌شوند.ابرپارامترها تأثیر به‌سزایی در سرعت و عملکرد شبکه دارند.
به طور کلی هر پارامتری را که مجبور شویم قبل از آموزش شبکه تعیین کنیم شامل ابرپارامترها می‌شود.

۲-٤ کاهش ویژگی

کاهش ویژگی به دو روش انتخاب ویژگی و استخراج ویژگی تقسیم می‌شود که در انتخاب ویژگی یکسری از ویژگی های کم اهمیت به طور کامل حذف می‌شوند اما در روش استخراج ویژگی ، ویژگی ها با هم ادغام و ویژگی جدیدی را که برای ما قابل فهم نیست تولید می‌کند.به برخی از دلایل کاهش ابعاد می‌توان به سرعت الگوریتم با داده‌هایی با ابعاد کم بیشتر ای و نیاز به فضای ذخیره‌سازی کمتری است و احتمال overfit را کاهش می‌دهد. روش تحلیل مولفه‌های اساسی(PCA) یکی از روش استخراج ویژگی که براساس تعیین مولفه‌های اساسی (PC) در داده‌ها کار می‌کند . مولفه های اساسی در حقیقت همان بردار ویژه‌های ماتریس کوواریانس داده ها هستند. بیشترین واریانس داده‌ها در راستایی قرار دارد که بردار ویژه‌ی متناظر با بزرگترین مقدار ویژه در آن راستا قرار دارد. به همین ترتیب هر چقدر مقدارویژه کوچکتر شود واریانس داده ها در راستای بردار ویژه متناظر با آن کمتر می‌شود.

۲-۵ مروری بر کارهای انجام شده

کارهای زیادی در خصوص سیستم‌های تشخیص نفوذ انجام شده است که در ادامه به بررسی برخی از مقالاتی که در زمینه یادگیری عمیق کار کرده‌اند و معیار دقت تشخیص نفوذ که یکی از چالش‌های پراهمیت در این حوزه که بسیار مورد توجه محققان قرارگرفته است می‌پردازیم.

در این تحقیق [20] از مفهوم انباشته‌شدن(تعمیم پشته) با استفاده از داده‌های ناهمگن برای شناسایی موثر نفوذ شبکه ارایه کرده‌اند. در این روش از دو مجموعه‌داده ناهمگن UNSSB-15 (شبیه سازی شده)و UGR16 (در زمان واقعی) ترکیب الگوریتم‌های جنگل تصادفی، رگرسیون لجستیک، نزدیکترین همسایه Kو ماشین بردار پشتیبانی منجر به پیش‌بینی‌های برتر باتوجه به یک مجموعه‌داده در زمان واقعی نسبت به یک مجموعه شبیه‌سازی شده است. اهداف اصلی محققان افزایش سرعت و دقت تشخیص نفود شبکه بوده است که نتایج مطلوبی حاصل شده است.

در تحقیقی دیگر[۶] ارایه یک روش تشخیص نفوذ با استفاده از یادگیری مجموعه‌داده محدود است. عملکرد تعبیه، عملکرد فاصله و عملکرد از دست دادن جنبه‌های اصلی است که بر نتایج طبقه‌بندی با استفاده از FSL[1] تاثیر می‌گذارد. در این روش از کمترین میزان داده برای آموزش استفاده شده است. از مجموعه‌داده NSL-KDD و UNSW-NB15 جهت تشخیص

[1] Few Shot Learning

نفوذ استفاده شده است، روش پیشنهادی در این تحقیق با الگوریتم‌هایی از جمله (RNN[1])، (RF[2])، (NB[3])، (SVM[4])، (MLP[5]) مقایسه شده است که برتری الگوریتم FSL نسبت به سایر الگوریتم‌ها اثبات شد.

در این پژوهش [21] یک مدل مجموعه‌ای را پیشنهاد می‌کنند که بهترین مدل‌های ML[6] و DL[7] را برای دستیابی به معیارهای عملکرد بالا ترکیب می‌کند.در نهایت بهترین مدل‌ها را با مجموعه‌داده CIC-IDS2017 مقایسه کردند و آنها را با مدل‌های مدرن مورد قبلی قرار دادند. در این تحقیق، یک معیار با عملکرد بالا با زمان آموزش نسبتا کم برای هر دو مل ML وDL در تشخیص نفوذ شبکه به دست آمد و CART[8] بهترین عملکرد را داشت ، مدل‌های ML به دلیل ابهام کمتر بردارهای ویژگی در مجموعه‌داده ، توانستند عملکرد خوبی داشته باشند. با مجموعه پیچیده‌تر ، مدل‌های DL ممکن است از مدل‌های ML بهتر عمل کنند. با این حال مدل‌های ML مزیت کنترل بر ویژگی دارند که در روش‌های DLوجود ندارد. اهمیت ویژگی به انتخاب بهترین ویژگی‌های مربوط به داده‌ها با معیار اطلاعات متقابل و کاهش مجموعه ویژگی‌ها برای دقت بهتر و محاسبات سریعتر در عملکرد کمک کنند.

در این پژوهش [۲۲] از الگوریتم شبکه عصبی در یادگیری عمیق جهت تشخیص نفوذ و افزایش امنیت در وسایل‌نقلیه مورد مطالعه قرارمی‌دهند. الگوریتم نزول گرادیان با تکانه و بهره تطبیقی(GDM/AG[9]) می‌تواند در مقایسه با الگوریتم شیب نزول با تکانه سنتی (GDM[10]) در تشخیص ناهنجاری خودرو سریعتر به همگرایی دست یابد و می‌تواند داده‌های ناهنجار را در میلی ثانیه تشخیص دهد . مدل پیشنهادی جیایان ژانگ می‌تواند خود را برای تشخیص حملات نا شناخته در زمان واقعی به میزان صحت ۹۷ تا ۹۸ درصد در جهت سازگاری با انواع حملات نا شناخته تطبیق دهد.

در پژوهشی دیگر [۲۳] در ابتدا یک بررسی از رویکردهای یادگیری عمیق برای تشخیص نفود امنیت سایبری و سپس ۳۵ مجموعه داده سایبری معروف را توصیف کرده‌اند و در هفت دسته مجوعه داه‌های مبتنی بر ترافیک شبکه، مجموعه داده‌های مبتنی بر شبکه الکتریکی، مجموعه داده‌های مبتنی بر ترافیک اینترنتی، مجموعه داده‌های مجازی مبتنی بر شبکه خصوصی، مجموعه داده‌های مبتنی بر برنامه‌های اندروید، مجموعه داده‌های مبتنی برترافیک اینترنت اشیا و مجموعه داده‌های مبتنی بر دستگاه‌های متصل به اینترنت طبقه‌بندی کرده‌اند و هفت مدل یادگیری عمیق شامل شبکه‌های عصبی مکرر، شبکه‌های عصبی عمیق، شبکه‌های عصبی متحرک، رمزگذاری خودکار عمیق، ماشین‌های محدود شده بولتزمن، ماشین‌های عمیق بولتزمن، شبکه‌های اعتقادی عمیق ارایه داده‌اند و عملکرد این هفت رویکرد را در دو مجموعه‌داده ترافیک واقعی شامل CSE-CIC-IDS2018 و BOT-IOT مورد بررسی قرار گرفته است و در نهایت با چهار رویکرد یادگیری ماشین یعنی بیز ساده، شبکه‌ی عصبی مصنوعی،جنگل تصادفی و ماشین‌بردار پشتیبان سنجش می‌کنند.

---

[1] Recurrent neural network
[2] Random Forest
[3] Naive Bayes
[4] Support vector machine
[5] Multi layer perceptron
[6] Machine learning
[7] Deep learning
[8] Classification and regression tree
[9] gradient descent momentum /adaptive gain
[10] gradient descent with momentum

در این تحقیق [۲۴] باتوجه به اهمیت امنیت اینترنت اشیا (IOT) یک استراتژی تشخیص حملات سبک با استفاده از یک ماشین‌بردار پشتیبان مبتنی بر یادگیری (SVM) برای تشخیص حملات DOS درنظر گرفته‌اند. تنها ویژگی در نظرگرفته شده در این مقاله میزان ورود بسته به گره است. برای طبقه‌بندی از یک طبقه‌بندی‌کننده SVM با ورودی در قالب دو یا سه ویژگی غیر مجتمع استفاده می‌شود و با سایر طبقه‌بندی کننده‌های مبتنی بر یادگیری ماشین از جمله: NN، K-NN، DT مقایسه شده است تا زمان تشخیص و دقت طبقه‌بندی با SVM را با ترکیبی از دو یا سه ویژگی پیچیده رضایت‌بخش را نشان دهد.

در این مقاله [۲۵]کاهش بعد و طبقه‌بندی دو لایه برای تشخیص حملات با فرکانس پایین در ستون فقرات IOT هدف محققان در این تحقیق تشخیص دقیق حملات(R2L)، (U2R) بدون کاهش عملکرد و پایین بودن نرخ مثبت کاذب به دلیل استفاده از ویژگی اصلاح و نرخ تشخیص کلی بالاتر به دلیل استقرار طبقه‌بندی‌کننده چندلایه و پیچیدگی محاسباتی کمتر به دلیل استقرار کاهش ابعاد در دو لایه بوده است و مجموعه‌داده مورداستفاده NSL-KDD بوده است. مدل پیشنهادی از مدل‌های مشابه موجود بهتر عمل کرده‌است و میزان تشخیص برای هر دو حملات فرکانس پایین و معمولی می‌باشد.

در مطالعه پژوهش[۲۶] یک چارچوب جدید یادگیری ماشین نیمه‌سطحی(MSML) برای سیستم تشخیص نفوذ ارایه داده‌اند که دو مشکل عدم تعادل ترافیک شبکه و توزیع غیر یکسان بین مجموعه آموزش و مجموعه آزمون که بر استحکام مدل خیلی تاثیر می‌گذارد. مدل MSML با بهبود الگوریتم K-MEANS الگوریتم نوین‌تری به نام HSK-MEANS ارایه‌داده‌است که برای یکسان‌سازی داده‌ها از نرمال‌سازی لگاریتمی استفاده می‌کند. اما در گزینش پارامترها منعطف نمی‌باشد. برای ارزیابی این مدل از مجموعه‌داده KDDCPU99 استفاده شده است. نتایج بدست‌آمده نشان می‌دهد که از نظر دقت کلی به میزان ۹۶.۶ درصد نسبت به مدل‌های دیگر تشخیص نفوذ کاراتر است.

در مقاله پیش‌رو [۱۱] سیستم تشخیص نفوذ شبکه مبتنی بر شبکه‌عصبی عمیق با واحدهای مکرر دروازه دار (GRU[1]) که از GRU به عنوان واحد حافظه اصلی ،همراه با MLP برای شناسایی نفوذهای شبکه استفاده می‌کند. آزمایش بر روی مجموعه داده‌های معروف KDD99 و NSL-KDD است .میزان تشخیص کلی ۹۹.۹۴ در KDD99 و ۹۹.۳۱ در NSL-KDD بود. آزمایشات مقایسه ای روش GRU در مقابل LSTM[2] نشان داده که سیستم دارای عملکرد پیشرو است.

در پژوهشی دیگر[۱۲] محققان در این تحقیق یک روش تشخیص نفوذ شبکه مبتنی بر یادگیری خودآموز براساس چارچوب STL[3] با ترکیب دو الگوریتم SVM , SAE[4] پیشنهاد کرده‌اند. مدل پیشنهادی با استفاده از مکانیزم خودکار رمزگذار که یک الگوریتم یادگیری موثر برای ویژگی‌های جدید و کاهش ابعاد به صورت بدون نظارت است و استفاده از SVM به جای SOFT-MAX برای طبقه‌بندی استفاده شده است که عملکرد خوبی در طبقه‌بندی باینری و چندکلاسه نسبت به روش‌های قبلی مانند بیز ساده، J48، RF، از خود نشان داده است و استفاده از مجموعه داده NSL-KDD برای آزمایش است.روش پیشنهادی STL-IDS باعث افزایش دقت طبقه‌بندی و سرعت آزمایش شده که تشخیص نفوذ شبکه را بهبود بخشیده است.

[1] Gated recurrent units
[2] Long short-term memory
[3] Self taught learning
[4] sparse auto encoder

در این تحقیق[۲۷] یک روش جدید کاهش ابعاد را پیشنهاد می کنند که در آن دو تکنیک انتخاب ویژگی و استخراج باهم ادغام شده است.در این روش رویکرد کسب اطلاعات (IG[1]) و تجزیه و تحلیل مولفه‌های اصلی (PCA[2]) برای کاهش ویژگی‌ها و استخراج مجموعه جدیدی از ویژگی‌ها غیرمرتبط استفاده شده است در حالی که طبقه‌بندی‌کننده گروهی بر اساس SVM، IBK[3]، MLP برای ساخت مدل طبقه‌بندی استفاده می شود. سپس از الگوریتم AOP برای به‌دست آوردن تصمیم نهایی برای تشخیص طبیعی بودن یا حمله بودن یک نمونه استفاده شده‌است عملکرد این روش IG-PCA بر اساس سه مجموعه‌داده معروف ISCX2012، NSL-KDD، Kyoto2006+ ارزیابی می‌شود. تجزیه و تحلیل مقایسه‌ای روش پیشنهادی نشان می‌دهد که عملکرد بهتری در مورد دقت طبقه‌بندی و میزان تشخیص و میزان هشدارهای کاذب نسبت به اکثر رویکردهای مدرن موجود دارد.

در این مقاله[13] برای تشخیص نفوذ شبکه از رمزگذار خودکار پراکنده (SSAE[4]) که نمونه‌ای از یک استراتژی یادگیری عمیق ، برای استخراج ویژگی‌های پراکنده با ابعاد بالا استفاده کرده‌اند .نتایج تجربی بر روی مجموعه‌داده NSL-KDD نشان می‌دهد که SSAE با ساختار بهینه می‌تواند ویژگی‌های اصلی را بدون ازدست دادن مقدار اطلاعات موجود بین داده‌های اصلی تا ۵ بعد فشرده کند.عیب روش پیشنهادی که محقق به ان اشاره کرده است این روش نمی‌تواند به طور موثر نمونه‌های حمله با فرکانس پایین U2R و R2L را تشخیص دهد یعنی نمی تواند بر عوارض ناشی از توزیع نامتعادل داده‌ها غلبه کند.

در مقاله‌ای دیگر[۵] رویکرد ترکیبی تشخیص نفوذ مبتنی بر شبکه عصبی با استفاده از ANN وFCM برای طبقه‌بندی الگوهای عادی و حمله ونوع حمله ارایه داده‌اند شبکه‌های عصبی مصنوعی از فرایندهای یادگیری که در سیستم‌های بیولوژیکی اتفاق می‌افتد الهام گرفته شده است سلول‌های عصبی سعی می کند مکانیسم‌های عملکرد بیولوژیکی خود را تقلید کنند. یادگیری را می‌توان به عنوان یک فرایند بهینه‌سازی درک کرد . از پرسپترون چندلایه (MLP) برای سیستم تشخیص نفوذ استفاده می‌شود نتایج نشان می‌دهد که سیستم پیاده‌سازی شده و طراحی شده حملات راشناسایی و در پنج گروه (NORMAL، DOS، U2R، R2L، PROBE) طبقه‌بندی می کند و از مجموعه‌داده KDD برای آموزش و ارزیابی طبقه‌بندی‌کننده ANN استفاده شده است. الگوی ورودی به پنج خوشه که نمایانگر پنج کلاس هستند تبدیل شده است . مقادیر مرکز و مقادیر شاخص برای هرخوشه با استفاده از خوشه‌بندی میانگین C فازی محاسبه شده است.

در این بررسی[۱۰] الگوریتم‌های شبکه عصبی مصنوعی (ANN) به عنوان طبقه‌بندی‌کننده برای تشخیص سوابق عادی و حمله مورد بررسی قرار داده‌اند و از مجموعه‌داده KDD CUP 99 برای محقق کردن رساله خود استفاده کرده‌اند در این تحقیق ثابت شده است که FFNN با ۱۰نورون و۲ لایه نسبت به FFNN با تعداد مختلف نورون و ۲ لایه بهتر عمل کرده است.

در بررسی دیگر[۲۸]یک IDS برای ترافیک شبکه بی سیم همرا با تکنیک استخراج ویژگی‌های مبتنی بر پوشش ارایه داده‌اند که روش استخراج بهینه ویژگی با استفاده از الگوریتم درختان اضافی(ET) پیاده‌سازی شده است.اثر بخشی و کارایی

[1] information gain
[2] principal component analysis
[3] Instance-based learning
[4] stacked sparse auto encoder

این روش با استفاده از مجموعه‌داده UNSW-NB15 ،AWID مورد مطالعه قرار گرفته است علاوه‌براین wfeu-ffdnn با الگوریتم‌های استاندار یادگیری ماشین شامل بیز ساده(NB)، جنگل تصادفی(RF)، ماشین بردار پشتیبان(SVM)، درخت تعمیم(DT) و نزدیکترین همسایه K (KNN) مقایسه می‌شود. مطالعات تجربی شامل انواع حملات دوتایی و چند کلاسی است. نتایج نشان داد که WFEU-FFDNN پیشنهادی دقت تشخیص بیشتری نسبت به سایر رو ش‌ها دارد.

در این مطالعه[۲۹] روش AdaBoost را برای بهبود عملکرد سیستم تشخیص نفوذ پیشنهاد داده‌اند که از تکنیک ابر نمونه‌گیری اقلیت مصنوعی(SMOTE) ، تجزیه و تحلیل اجزای اصلی (PCA) و انتخاب ویژگی‌های مجموعه (EFS) در مجموعه‌داده CIC-IDS-2017 ثابت شده است. مطالعات تجربی نشان می‌دهد روش پیشنهادی عملکردی بهتری از خود نشان می‌دهد.

در مطالعه‌ای دیگر[۳۰] با استفاده از روش سیستم تشخیص نفوذ جرقه یادگیری عمیق (DLS-IDS[1]) که شامل چهار قسمت اصلی می‌باشد انتخاب و کاوش، پیش‌پردازش، راه‌حل عدم تعادل کلاس و در نهایت آموزش برروی آپاچی جرقه است. در مجموعه داده NSL-KDD جهت افزایش دقت استفاده می کند که برای کاهش ویژگی مجموعه‌داده از روش ابر نمونه‌گیری اقلیت مصنوعی (SMOTE) به‌کار می‌برند که در نهایت دقت به ۸۳.۵۷ درصد بهبود می‌یابد.

در مقاله [۷] به معرفی یک الگوریتم تشخیص نفود ترکیبی مبتنی بر k-means، RF و یادگیری عمیق برای حل مسئله IDS پرداختند که از دو مجموعه داده NSL-KDD و CIC-IDS2017 برای ازیابی دقت تشخیص نفوذ استفاده کرده‌اند که به ترتیب شامل ۸۵.۲۴ درصد برای مجموعه داده NSL-KDD و ۹۹.۹۱ درصد برای مجموعه داده CIC-IDS2017 بدست آمده است. روش پیشنهادی سرعت پیش‌پردازش داده سریع‌تر و زمان آموزش بالقوه کمتری داشت.

در پژوهش[۸] یک IDS برپایه یادگیری عمیق به کمک شبکه‌های عصبی عمیق (FFDNN[2]) همراه با یک الگوریتم انتخاب ویژگی مبتنی بر فیلتر پیشنهاد می کنند. FFDNN-IDS با استفاده از مجموعه داده‌های شناخته شده NSL-KDD ارزیابی می کنند و با روش‌های یادگیری ماشین موردمقایسه قرار می‌دهند.که دقت تشخیص نفوذ در این الگوریتم۹۹.۳۷ درصد است. و نسبت به سایر روش‌های یادگیری ماشین برتری دارد.

در پژوهشی دیگر[۹] یک سیستم هشدار تشخیص نفوذ ترکیبی را با مدل یلدگیری عمیق و با شبکه‌های عصبی عمیق پیشنهاد کرده‌اند که یک ارزیابی جامع از آزمایش‌های DNN را برروی مجموعه‌داده KDDCUP 2017، NSL-KDD، UNSW-N B15، WSN-DS، KYTON، CICIDS 2017 صورت گرفته است که دقت تشخیص نفوذ به ترتیب به صورت NSL-KDD و KDDCUP99 در محدوده ۹۵ درصد تا ۹۹ درصدمی‌باشد و برای UNSW-NB15 و WSN-DS در محدوده ۶۵ درصد تا ۷۵ درصد است.

در این مقاله[۱] با استفاده ازروش ترکیبی تحلیل مولفه‌های اصلی (PCA) و یادگیردی عمیق(DL) روش (PCA-DL[3]) را ارایه داده‌اند که دقت تشخیص حملات در مجموعه‌داده NSL-KDD به میزان ۹۲ درصد رسیده است و با سایر الگوریتم از

[1] Deep Learning Spark-Intrusion Detection System
[2] feed forward deep neural networks
[3] Principle Component Analysis-Deep learning

قبیل جنگل تصادفی، درخت تصمیم، بیز ساده، رگرسیون لجستیک، شبکه‌های عصبی مصنوعی، الگوریتم نزدیکترین همسایه، الگوریتم ماشین بردار پشتیبان(SVM) مقایسه کرده‌اند و نسبت به دیگر الگوریتم‌ها از دقت بالاتری برخوردار است.

جدول ۲- ۱. مقایسه الگوریتم های مرور شده

ردیف	سال-تحقیق	روش انجام پژوهش	مجموعه داده
۱	۲۰۲۱-[۷]	الگوریتم ترکیبی مبتنی برK-MEANS توزیع‌شده، RF و یادگیری عمیق	NSL-KDD CIC-IDS2017
۲	۲۰۲۰-[۱]	الگوریتم PCA-DL: تجزیه‌وتحلیل مولفه‌های اصلی با یادگیری عمیق	NSL-KDD
۳	۲۰۲۰-[۶]	الگوریتم FSL: یادگیری مجموعه‌داده محدود	NSL-KDD UNSW-NB15
۴	۲۰۲۰-[۳۰]	الگوریتم DLS-IDS : تشخیص نفوذ با جرقه یادگیری عمیق	NSL-KDD
۵	۲۰۱۹-[۸]	الگوریتم FFDNN: مبتنی بر یادگیری عمیق و شبکه‌های عصبی عمیق	NSL-KDD
۶	۲۰۱۹-[۹]	الگوریتم ترکیبی با مدل یادگیری عمیق و شبکه‌های عصبی عمیق	NSL-KDD KDDCUP99 UNSW-NB15 WSN-DS
۷	۲۰۱۸-[۱۱]	الگوریتم GRU: مبتنی یر شبکه‌های عصبی عمیق با واحدهای مکرر	NSL-KDD KDD 99
۸	۲۰۱۸-[۱۲]	الگوریتم STL: مبتنی بر خودرمزگذار	NSL-KDD
۹	۲۰۱۸-[۱۳]	الگوریتم SSAE: رمزگذار خودکار پراکنده	NSL-KDD
۱۰	۲۰۱۸-[۲۷]	الگوریتم IG-PCA: رویکرد کسب اطلعات با تجزیه و تحلیل مولفه‌های اصلی	ISCX 2012 NSL-KDD +KYOTO 2006

۲-۶ جمع‌بندی

تکنیک‌ها و الگوریتم‌های متفاوتی برای توسعه یک سیستم تشخیص نفوذ برای شناسایی و طبقه‌بندی حملات سایبری استفاده می‌شود. با این وجود از آنجایی که حملات مخرب مداوم در حال تغییر هستند و در حجم عظیمی از داده‌ها رخ می‌دهد که نیازمند راه‌حلی مقیاس‌پذیر است چالش‌های زیادی به‌وجود می‌آید. مجموعه داده‌های بدافزار برای تحقیقات بیشتر توسط جامعه امنیت سایبری به صورت عمومی ارائه شده است. با این حال هیچ تجزیه و تحلیل دقیقی از عملکرد الگوریتم‌ها در مجموعه داده‌های مختلف در دسترس عموم قرار داده نشده است و با توجه به ماهیت پویایی بدافزارها با روش‌های حمله‌ایی که به صورت مداوم تغییر می‌کند نیازمند تحقیقات خواهد بود.

۳. روش پیشنهادی

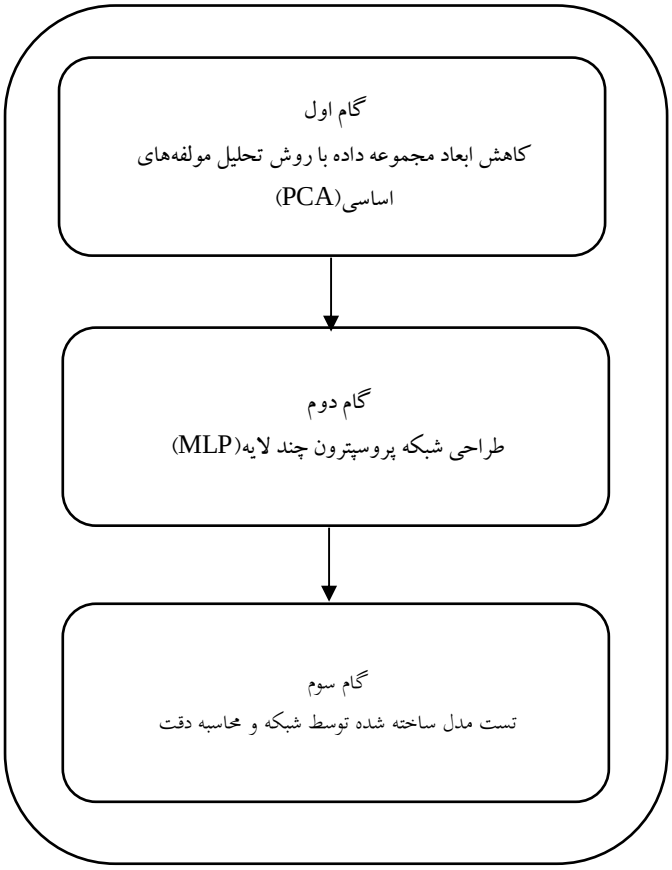
۳-۱ مقدمه

در فصل قبل درباره روش‌های افزایش دقت در سیستم‌های تشخیص نفوذ [۵]،[۷]، [۸]، [۱۱]، [۱۳] بحث کردیم. همانگونه که گفته شد در طی این چندسال تحقیقات بسیار زیادی در زمینه‌های مختلف برای افزایش دقت سیستم های تشخیص نفوذ انجام شده است. یکی از روش‌ها در این سیستم‌ها شبکه‌های پروسپترون چندلایه(MLP)، که یکی از روش‌های یادگیری عمیق است، برای حل چالش دقت ارائه شده است. افزایش تعدا لایه‌ها درmlp کارایی این شبکه را بهبود می‌بخشد. با اضافه کردن هر لایه به شبکه و به اصطلاح عمیق‌تر کردن شبکه که برای حل مسائل پیچیده‌تر مورد استفاده قرار می‌گیرد. اما از افزایش بار محاسباتی که متحمل می‌شود نباید غافل بود پس در انتخاب لایه با توجه به مساله باید دقت کافی به عمل آید.یکی از الگوریتم‌های مرسوم بهینه‌سازی وزن‌ها در MLP روش گرادیان کاهشی ریز دسته‌ای می‌باشد که یکی از معایب این الگوریتم یادگیری نرخ یادگیری می‌باشد که سبب کاهش همگرایی می‌شود و همچنین در طراحی شبکه برای هر لایه توابعی را در نظر می‌گیریم. تابع فعالساز جزء اصلی گره‌های شبکه‌های عصبی محسوب می‌شود. توابع فعالسازی در شبکه‌های عصبی تعیین می‌کند که ایا گره باید فعال شود یا غیرفعال باشد یکه از توابع فعالساز پرکاربرد تابع RELU می باشد که در کنار مزایای مثل از لحاظ محاسباتی کاراتر از توابع دیگر است و همینطور به دلیل ویژگی خطی بودن تابع، همگرایی الگوریتم گرادیان کاهشی را سریع‌تر می‌کند اما مشکلی که دارند مشکل مرگ نرون‌ها در حالتی که خروجی چندین تابع RELUبه طور متوالی برابر با صفر باشد.

بنابراین باید الگوریتم بهینه‌سازی وزن ها و تابع فعالساز به گونه‌ای انتخاب شود که مشکل نرخ یادگیری درگرادیان کاهشی و مرگ نرون‌ها در تابع فعالساز RELUرا حل کند و در نهایت سبب سرعت همگرایی و بهبود دقت در سیستم‌های تشخیص نفوذ شود. در ادامه این فصل یک نمای کلی از الگوریتم پیشنهادی را نمایش می‌دهیم. سپس مراحل الگوریتم پیشنهادی خود را توضیح داده و در نهایت مطالب گفته شده را جمع‌بندی می‌کنیم.

۳-۲ الگوریتم روش پیشنهادی

در این بخش یک نمای کلی از الگوریتم پیشنهادی خود را جهت بهبود افزایش دقت در سیستم‌های تشخیص نفوذ با روش PCA کار کاهش ویژگی را در مجوعه داده NSL-KDDاعمال می‌کنیم و با طراحی شبکه‌های عصبی پروسپترون چندلایه، که در آن از الگوریتم[۱۴] ADAM جهت بهینه‌سازی وزن ها و استفاده از تابع فعالساز RELU LEAKY [۱۴] بکار رفته است شرح می‌دهیم. همانطور که در شکل ۳-۱ مشاهده می‌کنید در ابتدای فاز اول از الگوریتم تحلیل مولفه‌های اساسی(PCA) جهت کاهش ابعاد مجموعه داده استفاده می‌کنیم. در فاز دوم از شبکه‌های پروپسترون چندلایه جهت تشخیص حملات ازداده‌های نرمال مورد استفاده قرار می‌گیرد و در آخرین گام یعنی گام سوم در مدل ساخته شده با داده‌های آزمایشی جهت بررسی عملکرد دقت در الگوریتم پیشنهادی اجرا می‌گیریم تا میزان دقت بررسی شود.

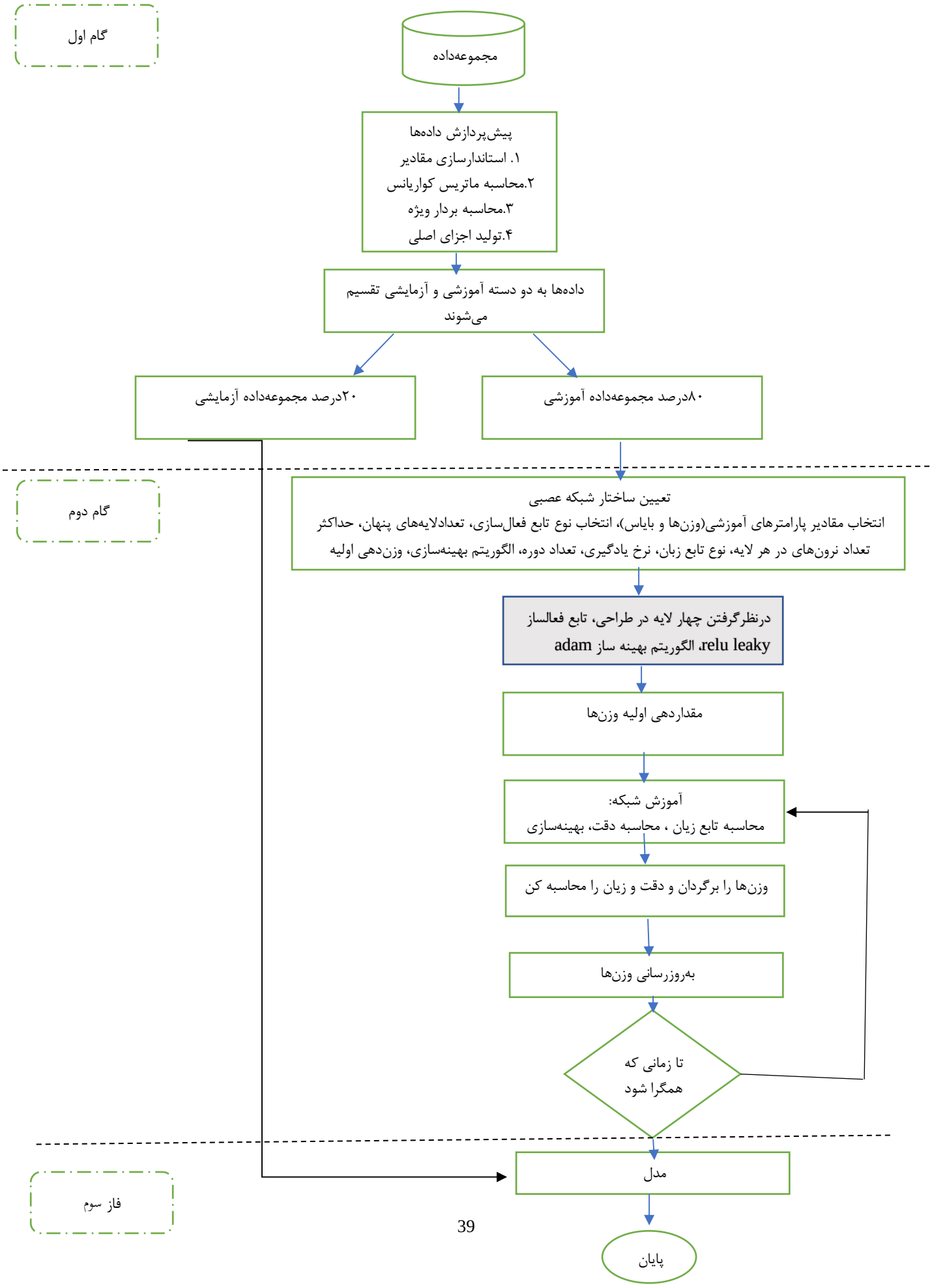


شکل ۳- ۱. نمای کلی الگوریتم پیشنهادی

۳-۳ جزئیات روش پیشنهادی

در این بخش ما برای بهبود دقت سیستم‌های تشخیص نفوذ از یکه از روش‌های یادگیری عمیق بهنام شبکه‌های پروسپترون چندلایه که در طراحی آن از تابع فعالساز LEAKY RELU [14] به جای RELU[۱] و الگوریتم بهینه[۱۴] ADAM به جای الگوریتم گرادیان کاهشی[۱] استفاده کنیم که به طور مفصل توضیح داده خواهد شد. در ادامه فصل در گام اول با الگوریتم تجزیه و تحلیل مولفه‌های اصلی(PCA[1]) کار کاهش بعد را برای افزایش بهبود سرعت پردازش داده‌ها صورت می‌گیرد که توضیح داده خواهد شد و در گام دوم از شبکه‌های عصبی عمیق پروستون چند لایه استفاده می شود تا نرمال یا حمله بودن رفتار شبکه مشخص شود

[1] Principal Component Analysis



شکل ۳- ۲. فلوچارت الگوریتم پیشنهادی

۳-۳-۱ گام اول

هدف حیاتی تحلیل و تجزیه مولفه‌های اساسی کاهش ویزگی یک مجموعه داده متشکل از تعداد زیادی متغییر مرتبط بایکدیگر است، در حالی که تاحد امکان تغییرات موجود در مجموعه داده را حفظ می‌کند. این امر با تبدیل به مجموعه جدیدی از متغییرهای مولفه اصلی(PC[1]) که همبستگی ندارند و مرتب شده‌اند به‌دست می‌آید که چند نفر اول بیشترین تغییرات موجود در همه متغییرهای اصلی را حفظ می‌کند.[۳۱]

تمرکز اصلی PCA تأکید بر هویت ها و واریانس ها در داده ها و شناخت الگوهای موجود در داده ها در مجموعه داده است. اساساً PCA در چهار مرحله کاهش ابعاد را انجام می دهد که در زیر به این چهار مرحله اشاره‌ایی می‌کنیم.

اولین قدم استاندارسازی مقادیر در ویژگی‌های مجموعه‌داده برای از بین بردن نتایج مغرضانه است. به عنوان مثال، در یک مجموعه‌داده می‌توان مقادیر دامنه بزرگتر (۰ تا ۱۰۰) را به عنوان دامنه کوچکتر (۰ تا ۱۰) کنترل کرد. با محاسبه میانگین و واریانس هر صفت حاصل خواهد شد. این ممکن‌است منجر به نتایج مغرضانه شود. برای جلوگیری از این PCA تمام مقادیر صفات را در یک مقیاس یکسان نشان می‌دهد.

دومین قدم ماتریس کوواریانس را محاسبه می‌کنیم PCA برای یافتن ویژگی‌های همبسته ماتریس کوواریانس را محاسبه می‌کند. ورودی‌های ماتریس کوواریانس مثبت یا منفی هستند. اگر مثبت باشد، ویژگی‌های مربوطه با یکدیگر در ارتباط هستند. این بدان معناست که هر دو ورودی می‌توانند برای هر پارامتری افزایش یا کاهش پیدا کنند. اگر مقدار ورودی منفی باشد ، ویژگی‌ها با یکدیگر رابطه معکوس دارند یعنی یک مقدار کاهش می‌یابد و مقدار دیگر افزایش می‌یابد.

سومین مرحله مقدار ویژه و بردارهای ویژه برای محاسبه واریانس بین ویژگی های مجموعه داده را از ماتریس کوواریانس بدست می‌آوریم. که بردارهای ویژه جهات را

چهارمین وآخرین مرحله تولید اجزای اصلی است. هر جزء اصلی مقدار واریانس صفات موجود در مجموعه‌داده را حمل می‌کند. حداکثر واریانس مجموعه‌داده در مؤلفه اصلی اول ذخیره می‌شود. حداکثر واریانس دوم در جزء اصلی دوم و غیره ذخیره می‌شود.هدف از این بخش بهبود سرعت پردازش داده هاست.[۱]

۳-۳-۲ گام دوم

در بخش دوم با طراحی شبکه با داده‌های آموزشی به شبکه آموزش حمله بودن یا نبودن داده‌ها را می‌دهیم.

الگوهای یادگیری عمیق با تحلیل پایدار داده‌ها و با پیدا کردن ساختارها و الگوها در داده‌ها آموزش را فرامی‌گیرند. فرایند آموزش با الگوهای محاسباتی به نام شبکه عصبی که از بنیاد مغز الهام گرفته شده، فراهم می‌شود. بنیاد این شبکه از تعدادی لایه پردازشی تشکیل شده است. در این بنیاد هرچه به لایه‌های بالاتر می‌رود قادر به حل مسائل سخت‌تری می‌شود. لایه ورودی داده‌های خام را پردازش و لایه‌های بعدی توانایی استفاده از اطلاعات نرون‌هایی که در لایه قبلی

[1] Principal Component

بدست آمده است را جهت به‌دست آوردن اطلاعات پیچیده‌تراز داده‌ها را فراهم می‌سازد ونتایج در لایه خروجی قابل مشاهده می‌باشد

فریضه‌ی شبکه‌ی عصبی، یادگیری میزان صحیح وزن‌ها و انحرافات در شبکه است به گونه‌ای که به‌طور مثال بتواند حمله و نرمال را از یکدیگر تفکیک کند. این کار با تکرار در شبکه‌های عصبی انجام می‌شود. این‌گونه است که چندین مرتبه داده‌هایی به الگوریتم داده می‌شود و در هر تکرار الگوریتم باید مقادیر بایاس و وزن‌ها را به‌هنگام‌سازی کند. نخستین بار که الگوریتم شبکه عصبی اجرا می‌شود، این الگوریتم یک دنباله از مقادیر وزن‌ها و انحراف را به بردارهای وزن‌ها و انحراف می‌دهد اینکار سبب می‌شود تا بردارها یک دنباله مقدار اولیه داشته باشند. در این صورت در هر بار اجرا، شبکه‌ی عصبی اشتباه خود را محاسبه می‌کند و با استفاده از آن مقادیر وزن‌ها و انحراف و خطا را به‌روزرسانی می‌کند. کار محاسبه خطا با تابع هزینه یا تابع ضرر است. در حقیقت شبکه عصبی با هر بار اجرا که شاهد مقدار ضرری می‌شود می‌آموزد که به چه میزان باید وزن ها و بایاس را به‌روزرسانی کند.

معماری شبکه عصبی پیشنهادی دارای چهار لایه که در لایه ورودی یک لایه ، دو لایه در لایه پنهان و در لایه خروجی هم دارای یک لایه می‌باشد. لایه ورودی اطلاعات را دریافت و لایه پنهان اطلاعات را پردازش و لایه خروجی نتایج را نشان می‌دهد.در ادامه بخش به توضیح فراپارامترهای که قبل از آموزش شبکه باید مشخص شوند می‌پردازیم.

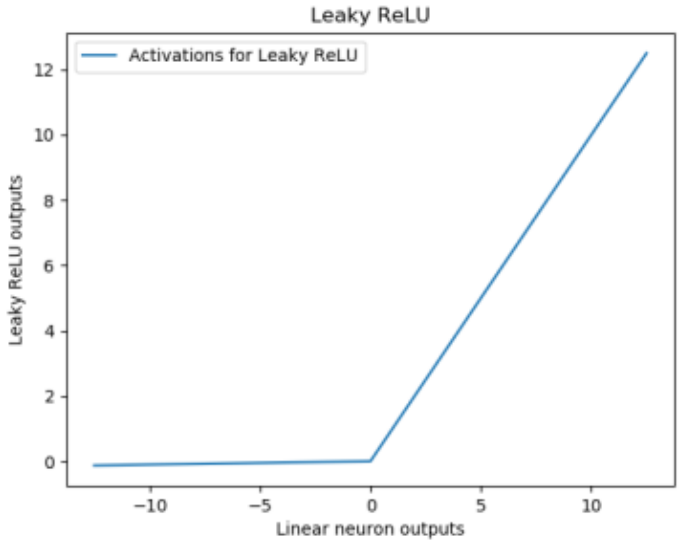
توابع فعالساز از جله فراپارامترهایی است که باید قبل از آموزش شبکه مشخص شود این توابه در تمام لایه های شبکه باید مشخص شود برای انتخاب توابع فعالساز پروسپترون چندلایه نمی‌توان از هر تابعی استفاده کرد بلکه باید حتما ماهیت مشتق‌پذیری، پیوسته، یکنوا نزولی را داشته باشد و همچنین مشتق اول آن به راحتی قابل محاسبه باشد.مانند RELU، LEAKY RELU می‌توان نام برد. (۳ – ۱)

تابع یکسوساز خطی رخنه‌دار(LeakyRelu): این تابع فعالسازی مشکل مرگ نرون‌ها در طول فرایند آموزش را ازبین می‌برد[۱۴]. برای افزایش سرعت مدل می‌توان از این تابع در لایه ورودی ودو لایه پنهان استفاده کرد.

$$LeakyRelu = \begin{cases} x & x \geq \text{صفر :اگر} \\ ax & x < \text{صفر :اگر} \end{cases}$$

(۱_۳)

شمایی از این تابع در شکل ۳-۳ قابل مشاهده است



شکل ۳- ۳. تابع فعالساز یکسوساز خطی رخنه دار [۱٤]

مزایا:

- مشکل مرگ RELU رارفع می کند.چرا که هنگام محاسبه مشتق اجازه یک گرادیان کوچک را می‌دهد.
- از لحاظ محاسباتی سریعتر عمل می کند.[۱۴]

تابع SOFTMAX این تابع که از آن در لایه خروجی استفاده می‌شود، تعمیم یافته‌های از تابع فعالساز سیگموئید بوده و برای حل مشکلات مربوط به دسته‌بندی از آن استفاده می‌شود. این امکان را فراهم می‌سازد که یک پیش‌بینی احتمالاتی را برای مسئله دسته‌بندی بیش از دو دسته انجام دهد و به صورت معادله (۳-۲) تعریف می شود.

$$S(y_i) = \frac{e^{y_i}}{\sum_j e^{y_i}}$$

(۳-۲)

تابع زیان از جمله فراپارامترهای می‌باشد که با توجه به ماهیت مسئله قابل انتخاب می‌باشد. برای محاسبه تابع زیان از تابع آنتروپی متقاطع که در مسائل گسسته دسته بندی استفاده می‌شود این تابع فاصله بین دو احتمال را محاسبه می کند و به صورت معادله (۳-۳) تعریف می‌شود.[۱۴]

$$cross\ Entropy\ (y, \hat{y}) = \frac{1}{n}\sum_{i=1}^{n}(y_i \log(\hat{y}))$$

(۳-۳)

y : مقدار واقعی مسئله

$\hat{y}$:خروجی تخمین زده شده از مسئله

y_i : خروجی هر نورون

مقداردهی اولیه یکنواخت گلوروت: مقداردهی اولیه یکنواخت گلوروت که همچنین با نام مقداردهی اولیه خاویر[۲] هم شناخته می‌شود، وزن‌های لایه L را براساس توزیع با میانگین صفر و انحراف معیار خاص در شبکه، به صورت یکنواخت در بازه:

$$\left[\sqrt{\frac{6}{(n_{l-1}+\ n_l)}}, \sqrt{\frac{6}{(n_{l-1}+n_l)}}\right]$$

(۴- ۳)

مقداردهی می‌کند. که در این معادله n_{l-1}وn_l تعداد نرون‌ها در لایه L-1 و L می‌باشد[۱۴]

الگوریتم بهین‌سازی و به‌روزرسانی وزن‌ها: بهین‌سازها الگوریتم‌هایی هستند که به واسطه به‌روزرسانی وزن‌ها در شبکه تلاش در جهت کم کردن تابع ضرر دارد. برای بهینه کردن مقدار وزن‌ها و بایاس در این پژوهش از الگوریتم ADAM استفاده می‌شود.

برآورد تکانه تطبیقی یا به اختصار (ADAM)، این روشی جهت محاسبه نرخ یادگیری تطبیقی برای هر پارامتر می‌باشد. این الگوریتم از مزایای الگوریتم‌های Adagrad و RMSprop استفاده می‌کند و میانگین فروپاشی نمایی از گرادیان‌های گذشته را در v_t ذخیره می‌کند. علاوه بر این ADAM، میانگین تکان‌های دوم گرادیان را در m_t ذخیره می‌کند.

m_t وv_t به ترتیب مقادیر میانگین (۳-۵) و واریانس (۳-۶) غیرمتمرکز هستند[۱۴]

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t$$

(۵_۳)

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_{t^2}$$

(۶_۳)

ADAM میانگین‌های تحرک نمایی گرادیان و گرادیان مربع را توسط معادلات (۳-۷) و (۳-۸) کنترل می‌کند[۱۴]

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t}$$
$$\hat{v}_t = \frac{v_t}{1-\beta_2^t}$$

که ابرپامترهایی با مقادیر $\beta_1 \in [0,1]$,β_2 است. معادله نهایی به‌هنگام‌سازی به صورت معادله (۳-۹)می‌باشد[۱۴]

$$w_{t+1} = w_t - \frac{\eta}{\sqrt{\hat{v}_t}} \odot \hat{m}_t$$

w_{t+1}: وزن جدید

w_t: وزن قبلی

η: نرخ یادگیری

$\hat{v}_t$: گرادیان مربع

$\hat{m}_t$: تحرک نمایی گرادیان

الگوریتم ADAMاز سایر روش‌های تطبیقی بهتر عمل می‌کند و خیلی سریع همگرا می‌شود. همچنین بر سایر مشکلاتی که الگوریتم‌های بهینه‌سازی همانند فروپاشی نرخ یادگیری، واریانس بالا در به‌هنگام‌سازی و همگرایی آهسته غلبه کرده‌اند، غلبه می‌کند[۱۴].

۳-۳-۳ گام سوم

در این بخش مدل‌سازی شبکه عصبی صورت گرفته است و با مجموعه داده آزمایشی مدل رامورد آزمایش قرار می‌هیم و معیار پر چالش دقت را بدست می‌آوریم.

۳-٤ جمع‌بندی

در این فصل ما الگوریتم پیشنهادی خود را شرح دادیم. همانطور که گفته شد الگوریتم تجزیه و تحلیل مولفه اساسی با یادگیری عمیق(PCA-DL) [۱] با استفاده از تحلیل مولفه اساسی ویژگی‌های مجموعه داده را کاهش می‌دهد و با شبکه پروسپترون چندلایه (MLP) جهت تشخیص فعالیت نرمال از حمله مورد استفاده قرار می‌گیرد که با توجه به دقت بدست آمده ما الگوریتم پیشنهادی خود را که در طراحی آن تعداد لایه‌ها را افزایش داده‌ایم و از الگوریتم بهینه ADAM جهت به‌روزرسانی استفاده‌کرده‌ایم و در لایه اول و دوم و سوم از طراحی شبکه از تابع فعالساز RELU.LAEKY جهت جلوگیری از مرگ نرون‌ها استفاده شده است.بنابراین ما در گام اول کار استاندارسازی مجموعه داده انجام دادیم و مجموعه‌داده را به دو قسمت جهت آموزش و آزمایش تقسیم کردیم در گام دوم کار طراحی شبکه را با درنظر گرفتن چهار لایه یکی در لایه ورودی و دوتا در لایه خروجی و یکی در لایه خروجی و مشخص کردن ابر پارامترها مثل تعداد نرون در هر لایه، الگوریتم بهینه سازی وزن‌ها، تابع زیان و توابع فعالساز در هر لایه انجام می‌دهیم و با هشتاد درصد مجموعه داده شبکه را مورد آموزش قرار داده تا مقادیر پارامترها از قبیل وزن نرون ها و بایاس را در این مرحله به

صورت بهینه مشخص شود و در انتها در گام سوم با بیست درصد دیگر از مجموعه داده شبکه مدلسازی شده را مورد آزمون قرار می‌دهیم تا دقت روش پیشنهادی در این مرحله مشخص شود

٤.روش تحلیل و ارزیابی

٤-١ مقدمه

دراین فصل قصد داریم تا الگوریتم پیشنهادی خود را شبیه سازی کرده و با الگوریتمPCA-DL[١]مورد مقایسه قرار دهیم و نتایج حاصل شده از روش خود را مورد بررسی قرار دهیم برای این‌کار ابتدا در محیط شبیه‌ساز python براساس مجموعه‌داده[1]NSL-KDD با تنظیم فراپارامترها الگوریتم پیشنهادی خود را طراحی می‌کنیم پس از پیاده‌سازی الگوریتم را از نظر معیارهای دقت ، صحت، خطا و یادآوری مورد بررسی قرارمی‌دهیم و در انتها ازلحاظ دقت با الگوریتم [1]PCA-DL مورد مقایسه قرار می‌دهیم.

٤-٢ روش ارزیابی

مجموعه‌داده [1]NSL-KDD که شامل ۴۱ ویژگی و۱۶۰۳۶۷ رکورد است. یک مجموعه داده ایست که به‌منظور رفع تعدادی از مشکلات مجموعه‌داده KDD99 مطرح شده است که در [۳۲]ذکر شده است. این مجموعه‌داده شامل پنج کلاس است که یک کلاس نرمال و چهار کلاس حمله DOS[1],R2L[2],U2R[3],PROB[4] است. که تعداد رکورد ها در جدول ۴-۱ مشخص شده است.

جدول ٤- ١. توزیع انواع حملات در مجموعه داده NSL-KDD [۳۳]

چند کلاسه					دسته‌بندی داده‌ها
U2R	R2L	Probe	Dos	نرمال	
۵۲	۹۹۲	۱۱۶۵۶	۴۵۹۲۷	۶۷۳۴۳	آموزشی
۲۰۰	۲۷۵۴	۲۴۲۱	۷۴۵۸	۹۷۱۱	آزمایشی
۲۵۲	۳۷۴۶	۱۴۰۷۷	۵۳۳۸۵	۷۷۰۵۴	کل

در شکل ۴-۱ نمونه رکورد از مجموعه‌داده NSL-KDD نمایش داده شده است.که باید دقت داشت سه ویژگی در این مجموعه داده به صورت عددی نیست و برای قابل فهم بودن شبکه باید تبدیل به عددی شود

[1] Denial of Service
[2] Remote to User
[3] User to Root
[4] Probing

```
0,tcp,ftp_data,SF,491,0,0,2,2,0.00,25,0.17,0.03,0.1
0,udp,other,SF,146,0,0,13,1,0.00,0,0.00,0.60,0.88,0
0,tcp,private,S0,0,0,0,123,6,1.00,26,0.10,0.05,0.00
0,tcp,http,SF,232,8153,0,5,5,0.20,55,1.00,0.00,0.03
0,tcp,http,SF,199,420,0,30,32,0.00,255,1.00,0.00,0.
0,tcp,private,REJ,0,0,0,121,19,0.05,19,0.07,0.07,0.
0,tcp,private,S0,0,0,0,166,9,1.00,9,0.04,0.05,0.00,
0,tcp,private,S0,0,0,0,117,16,1.00,15,0.06,0.07,0.0
0,tcp,remote_job,S0,0,0,0,270,23,1255,23,0.09,0.05,
```

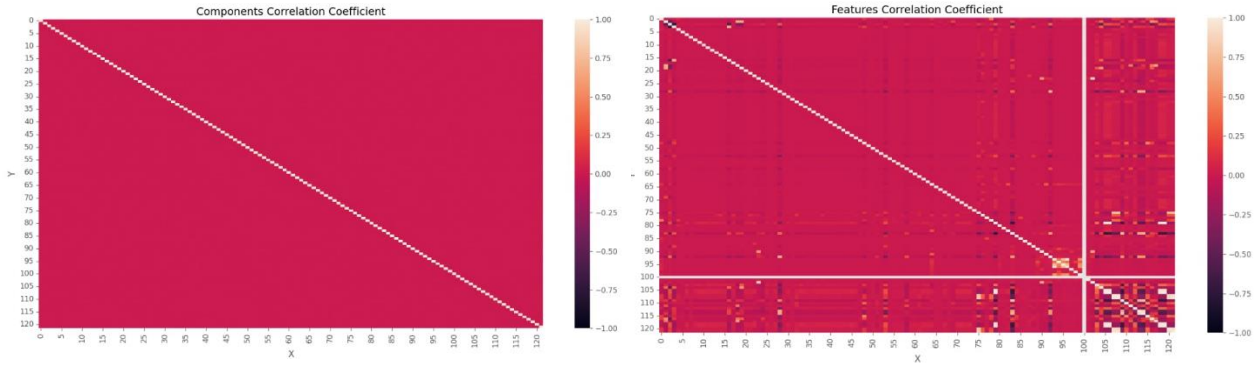
شکل ٤- ١. نمونه رکورد از مجموعه داده NSL-KDD [٣٤]

فرایند اعمال [١] PCA که به طور کامل در قسمت ۳-۳-۱ شرح داده‌شد به طور خلاصه می‌توان گفت طی فرآیند کاهش ابعاد در مجموعه داده NSL-KDD[١] می‌توان گفت روش PCA مواردی را به دنبال دارد از جمله:

- چگونگی ارتباط هر یک از متغییرها با یگدیگر به وسیله ایجاد ماتریس کوواریانس می‌توان پی برد.
- داده‌ها در چه جهاتی پراکنده هستند از طریق بردار ویژگی بدست می‌آید.
- مقدار اهمیت هریک از جهات‌ها که با مقادیر ویژه مشخص می‌شود

در نهایت این روش با ادغام ویژگی‌ها به ما اجازه می‌هد بردار ویژه‌هایی که نسبتا کم اهمیت هستند را حذف کنیم.

در شکل ۴- ۲ الف همبستگی بین ویژگی های اولیه مجموعه‌داده قبل از اعمال pca می‌باشد همانطور که میبینید شامل ۱۲۲ ویژگی است که همبستگی مناسبی بین ویژگی ها دیده نمی‌شود و درشکل۴-۲ ب میزان همبستگی بین ویژگی‌های اعمال شده با روش pca نشان می دهد درصورتی که تنها با انتخاب یک ویژگی می‌توان پوشش‌دهی خوبی روی مجموعه داده داشته باشیم.



الف) بدون اعمال pca ب) با اعمال pca

شکل ٤- ٢. نمودار همبستگی بین ویژگی های مجموعه داده

برای سنجش میزان کارایی مدل می‌توان مجموعه‌داده را به سه قسمت داده‌های آموزشی و داده‌های اعتبارسنجی و داده‌های آزمایشی تقسیم کنیم. که ابتدا آموزش را در بخش بزرگی از داده‌ها انجام می‌دهیم و سپس برای سنجش میزان کارایی مدل و قابلیت تعمیم دهی آن از داده‌های آزمایشی استفاده می‌کنیم. تعمیم‌دهی نشان‌دهنده‌ی میزان عملکرد مدل در برخورد با داده‌هایی است که تاکنون مدل آنها را در فرایند آموزش مشاهده نکرده است.

جهت اجرای دقیق شبکه در محیط برنامه‌نویسی پایتون با پردازنده xeon2core و 8 ROM صورت می‌گیرد و تنظیمات ابرپارامترها در شبکه عصبی عمیق به این صورت است که ازچهار لایه که شامل یک لایه ورودی و دو لایه پنهان و یک لایه آخر تشکیل شده است تعداد نرون‌ها در لایه ورودی همان تعداد ویژگی‌ها در شبکه است و به صورت برداری به شبکه تزریق می‌شود تعداد نرون‌ها در لایه پنهان اول ۸۶ و در لایه دوم ۲۵۶ و در لایه سوم ۱۲۸ در لایه آخر یعنی لایه خروجی شامل ۵ نرون می‌باشد.اندازه دسته ۳۲ در نظر گرفته‌ایم و همه‌ی لایه‌ها تماما بایکدیگر متصل هستند.تابع فعالساز برای لایه ورودی و دو لایه پردازش LeakyRelu و برای لایه خروجی softmax است. از الگوریتم بهینه‌ساز ADAM برای پیدا کردن وزن‌ها و بایاس بهینه استفاده می‌کنیم. در جدول ۴-۲ تنظیم مقادیر فراپارامترها درج شده است

در جهت تعیین بهینه‌ترین مقادیر فراپارامترهای جهت رسیدن به بهترین دقت از انواع روش‌های مقداردهی اولیه در مدل از قبیل تصادفی، انتقالی، یکنواخت Glorot و He که بهترین مقادیر از نظر دقت و بار محاسباتی مربوط به روش he و یکنواختGlorot می‌باشد و اندازه‌های دسته متفاوت و همین طور دوره و نرخ یادگیری مختلف استفاده شد و با آزمون و خطا نسبت به تعداد لایه ها و تعداد نرون‌ها تست و اجرا از شبکه گرفتیم که در هر کدام بهینه ترین‌ها را انتخاب و در جدول درج کردیم.

جدول ٤- ٢. تنظیم ابرپارامترها در شبکه mlp

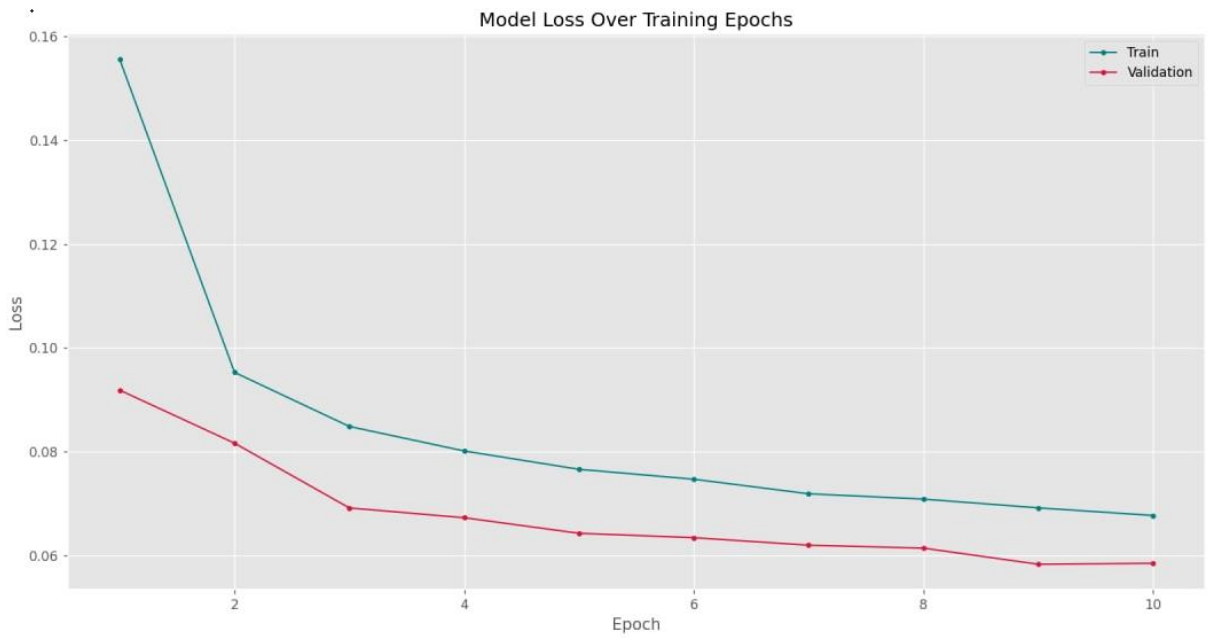
مقدار	نوع ابرپارامتر
۴	تعداد لایه
۸۶	تعداد نرون در لایه اول
۲۵۶	تعداد نرون در لایه دوم
۱۲۸	تعداد نرون در لایه سوم
۵	تعداد نرون در لایه چهارم
Leaky relu,softmax	توابع فعالساز
cross entropy	تابع زیان
adam	الگوریتم بهینه‌سازی
مقداردهی اولیه یکنواخت گلوروت	وزن‌دهی اولیه
۱۰	دوره
۶۴	اندازه دسته
۰.۰۰۰۲	نرخ یادگیری

٤-٣ تحلیل ارزیابی

در زیر به تعدادی از معیارها که با آن‌ها می‌توان عملکرد الگوریتم‌ها را مورد بررسی قرار داد اشاره و نتایج بدست آمده حاصل از مدلسازی روش پیشنهادی مورد تحلیل قرار می‌گیرد

٤-٣-١ خطا(Loos)

به اختلاف بین مقدار بدست‌آمده از مدل و مقدار واقعی خطا می‌گوییم همانطور که در نمودار ۴-۱ مشاهده می‌کنید خطا در این مدل به بعد از طی کردن ۱۰ دوره به میزان قابل توجهی کاهش یافته‌است



نمودار٤- ١. خطا در طول آموزش مدل

٤-٣-٢ صحت(Accuracy):

تعداد پیش‌بینی‌های درست بازگردانده‌شده توسط مدل را می‌توان با صحت عنوان کرد.

به صورت معادله در **٤-١** نشان می‌دهیم و با نمودار ۴-۲ می‌توان نحوه عملکرد و بهبود آن را در ۱۰ دوره مشاهده کرد.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

(۴-۱)

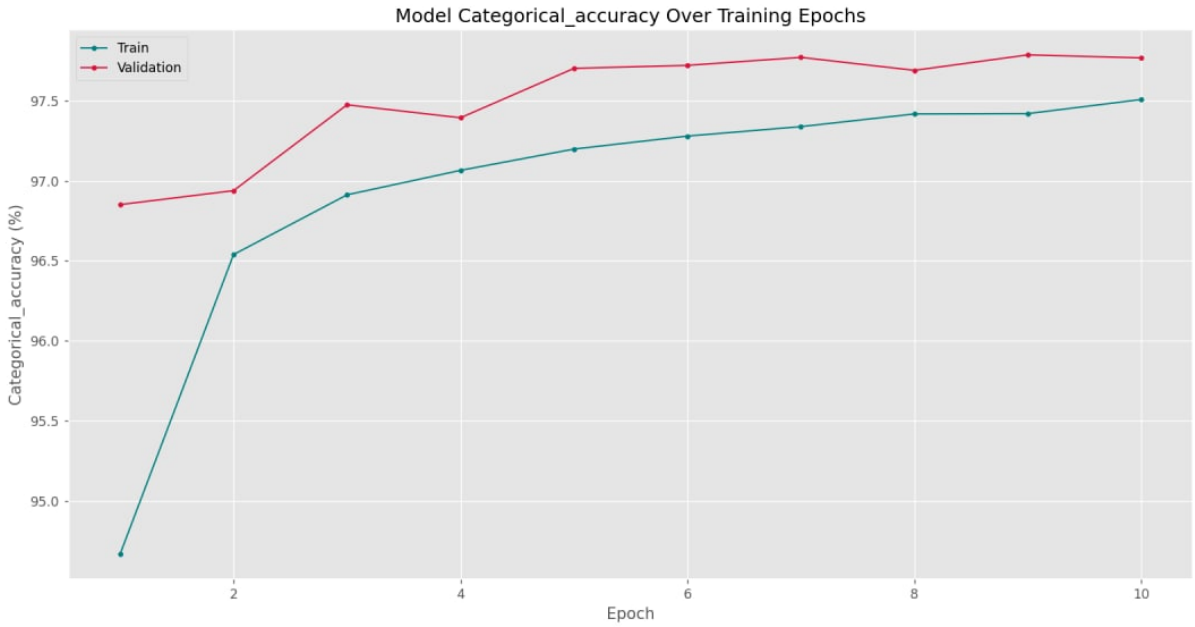
TP[1]: مثبت واقعی(پیش‌بینی‌های درست)

FP[2]: مثبت کاذب(مواردی که مدل به غلط پیش‌بینی کرده است)

FN[3]: منفی کاذب(مواردی که انتظار پیش‌بینی داشته ایم اما مدل پیش‌بینی نکرده است)

TN[4]: منفی واقعی(پیش‌بینی‌های نادرست)[۳۴]

[1] True Positive
[2] False Positive
[3] False Negative
[4] True Negative



نمودار٤- ٢. صحت در طول آموزش مدل

٤-٣-٣ یادآوری(Recall)

تعداد مثبت‌های برگردانده‌شده توسط مدل را می‌توان با عنوان یادآور(recall) یا حساسیت(sensitivity) نام برد که فرمول آن به صورت زیر می‌باشد

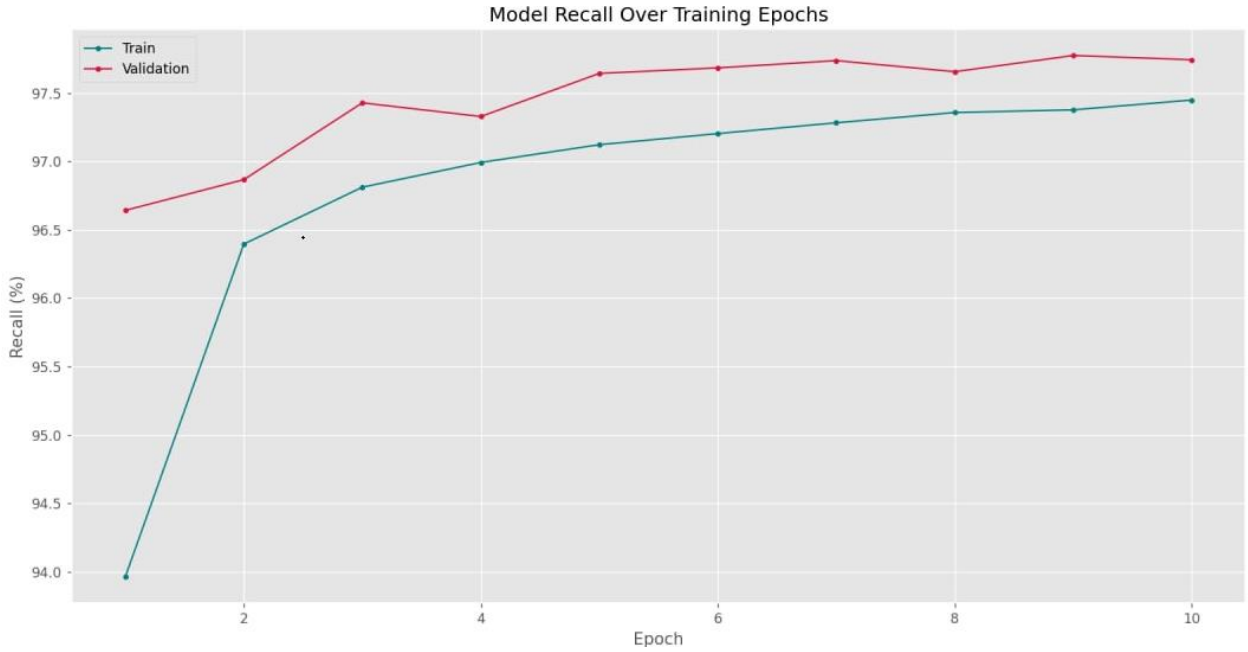
$$Recall = \frac{TP}{TP + FN}$$

(۲-۴)

TP: مثبت واقعی(پیش‌بینی‌های درست)

FN: منفی کاذب(مواردی که انتظار پیش‌بینی داشته ایم اما مدل پیش‌بینی نکرده است)[۳۴]

همانطور که در شکل۴-۳ مشاهده می کنید



نمودار٤- ۳. معیار یادآوری در طول آموزش مدل

٤-۳-٤ دقت (Precisions)

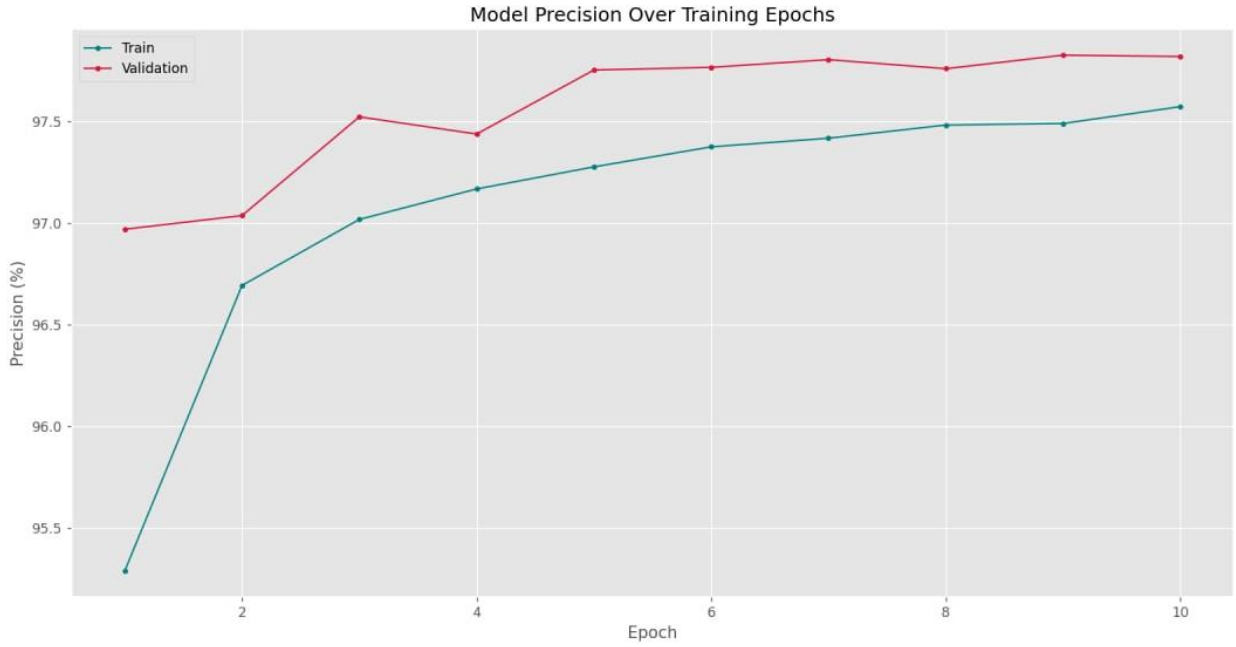
تعداد مقادیر صحیح برگردانده‌شده توسط مدل را می‌توان دقت نام برد. در واقع نرخ تشخیص هم می‌توان گفت که به صورت فرمولی زیر نمایش داده‌می‌شود:

$$Precisions = \frac{TP}{TP + FP} \tag{3-4}$$

TP: مثبت واقعی(پیش‌بینی‌های درست)

FP: مثبت کاذب(مواردی که مدل به غلط پیش‌بینی کرده است[۳۴]

همانگونه که در شکل ۴-۴می‌بینید دقت



نمودار٤- ٤. نمودار دقت در طول آموزش مدل

٤-٣-٦ گزارش طبقه‌بندی

این گزارش براساس امتیاز دقت(percision)، یادآوری(recall)، پشتیبان (support) و F1-score است

پشتیبان را میتوان اینگونه تفسیر کرد در قالب تعداد نمونه‌های پاسخ صحیح که در هر کلاس از مقادیر هدف قرار می‌گیرد.

F1-SCORE :این معیار میانگین بین دو مقدار یادآوری و دقت را در نظر می‌گیرد که در بهترین حالا یک و بدترین حالت صفر می‌باشد و معیاری مناسب جهت ارزیابی دقت یک آزمایش است.

در جدول ۴-۲ و۴-۳ و ۴-۴ گزارش طبقه‌بندی مربوط به مجموعه داده آموزش و اعتبار سنجی و آزمایش می‌باشد که به تفکیک برای هر کلاس معیارهای مختلف ارزیابی شده است که میانگین دقت (macro avg) برای مجموعه‌داده آموزشی و اعتبارسنجی و آزمایشی به ترتیب ۰.۹۴ درصد و ۰.۹۴درصد و ۰.۹۳درصد می‌باشد و همینطور معیار F1-score نیز برای سه مجموعه داده آموزش و اعتبار سنجی و آزمایش شامل ۰.۹۳ درصد است..

جدول ٤- ٣. گزارش طبقه بندی برای مجموعه داده آموزش

Model Classification Report on Train Dataset				
	percision	recall	F1-score	support
0	0.98	0.98	0.98	47395
1	0.99	1.00	1.00	34870
2	0.87	0.83	0.85	3794
3	0.89	0.81	0.85	272
4	0.97	0.98	0.97	9888
accuracy			0.98	96219
Macro avg	0.94	0.92	0.93	96219
Weghted avg	0.98	0.98	0.98	96219

جدول ٤- ٤. گزارش طبقه بندی برای مجموعه داده اعتبارسنجی

Model Classification Report on Validatin Dataset				
	percision	recall	F1-score	support
0	0.98	0.98	0.98	15963
1	0.99	1.00	0.99	11299
2	0.87	0.82	0.84	1399
3	0.90	0.79	0.84	87
4	0.97	0.98	0.97	3326
accuracy			0.98	32074
Macro avg	0.94	0.91	0.93	32074
Weghted avg	0.98	0.98	0.98	32074

جدول ٤- ٥. گزارش طبقه بندی برای مجموعه داده آزمایش

Model Classification Report on Test Dataset				
	percision	recall	F1-score	support
0	0.98	0.98	0.98	15848
1	0.99	1.00	1.00	11558
2	0.90	0.83	0.87	1310
3	0.83	0.82	0.82	93
4	0.96	0.97	0.97	3265
accuracy			0.98	32074
Macro avg	0.93	0.92	0.93	32074
Weghted avg	0.98	0.98	0.98	32074

٤-٤ مقایسه دقت

دقت مورد مقایسه در روش پیشنهادی ما با پروسپترون ۴ لایه و روش pca-dl [۱] شامل Macro avg می‌باشد که در جدول ۴-۵ قابل مشاهده می‌باشد

Macro avg: برای محاسبه این معیار ابتدا مقدار recall یا percision را بدست می‌آوریم در نهایت میانگین آن را محاسبه کرد:

$$macro\ average\ recall = \frac{\sum_{i=1}^{n} recall_i}{n} \quad (4\text{-}4)$$

$$macro\ average\ percision = \frac{\sum_{i=1}^{n} percision_i}{n} \quad (4\text{-}5)$$

جدول ٤- ٦. مقایسه دقت الگوریتم pca-dl و الگوریتم پیشنهادی

نام الگوریتم	دقت
الگوریتمpca-dl	۹۲ درصد
الگوریتم پیشنهادی	۹۳ درصد

٤-٤جمع‌بندی

ما در این فصل مدل پیشنهادی خود را در محیط برنامه‌نویسی python شبیه‌سازی کردیم سپس نتایج شبیه سازی را مورد نقد قراردادیم و با الگوریتم pca-dl [۱]مورد مقایسه قرار دادیم.با توجه به نتایج بدست آمده ما توانستیم دقت را به ۹۳ درصد افزایش دهیم.تعداد نرون‌ها را می‌توان گفت در بدست آوردن مقدار بهبنه دقت فراپارامتر باهمیتی است.در بخش بعدی به جمع‌بندی و همچنین پیشنهاد کار آینده می‌پردازیم.

۵. نتیجه‌گیری و پیشنهاد کار آینده

۵-۱ نتیجه‌گیری

در این پایان‌نامه مسئله بهبود دقت در سیستم‌های تشخیص نفوذ مطرح شده است. با توجه به افزایش استفاده از شبکه‌ها، امنیت اطلاعات در حین رد و بدل کردن آن‌ها بسیارحائز اهمیت می‌باشد روش‌های مختلفی جهت شناسایی، تشخیص و جلوگیری ارائه شده است اما با توجه به اینکه راه‌های نفوذ در حال تغییر است و هروزه هکرها در پی حملات جدیدی به سیستم‌ها با اهداف مختلفی هستند این حوزه نیازمند فعالیت و به‌روزرسانی می‌باشد.در فصل دوم به تعدادی از این روش‌ها از جمله الگوریتم‌های یادگیری عمیق [۱۲] ,[۱۳] ,[۶], [۱۱] **و** روش‌هاتی ترکیبی[۷] [۲۷][۱۱] توضیح داده‌شد. در اکثر این روش‌ها تلاش شده تا دقت تشخیص نفوذ بهبود یابد یکی از کارای انجام شده در این زمینه[۱] می‌باشد که از روش ترکیبی pca-dl کاهش بعد والگوریبتم یادگیری عمیق پروسپترون چندلایه جهت افزایش دقت استفاده می کنیم این روش به این صورت کار می کند که ابتدا با pca یک پیش‌پردازش روی مجموعه داده nsl-kdd اعمال می کنیم و سپس با الگوریتم پرسپترون چندلایه فعالیت های نرمال و حمله را تشخیص می‌دهد.

در این پایان‌نامه جهت حل مسئله بهبود دقت تشخیص نفوذ از روش تحلیل مولفه‌های اساسی و در طراحی شبکه mlp تعداد لایه‌ها را جهت بهبود عملکر شبکه افزایش دادیم و با چهار لایه طراحی صورت گرفت که در طراحی شبکه از الگوریتم adam جهت بهینه‌سازی وزن‌ها و تابع‌فعاساز leakly relu جهت جلوگیری از مرگ نرون انجام داده‌ایم.

ما روش پیشنهادی را به همراه تحلیل مولفه‌های اساسی و شبکه پروسپترون چهار لایه در محیط برنامه‌نویسی python اجرا کرده ایم با مقایسه نتایج پیاده‌سازی می‌توان دریافت که روش پیشنهادی با افزایش لایه‌های پنهان که کار پردازش را انجام‌می‌دهد توانستیم دقت تشخیص نفوذ را به ۹۳ درصد افرایش دهیم .این بهبود دقت نشان می‌دهد که روش پیشنهادی در مقایسه با روش [۱]عملکرد بهتری در رسیدن به افزایش دقت سیستم‌های تشخیص نفوذ داشته‌است.

۵-۲ پیشنهاد کار آینده

از آنجایی که روش پیشنهادی ما کاهش ویژگی مجموعه داده nsl-kdd را از طریق تحلیل مولفه‌های اساسی pca استخراج ویژگی صورت می گرفت این روش محدودیت‌هایی دارد از جمله تفسیرپذیری پایین اجزای اصلی در این روش و اجزای اصلی ترکیبی خطی از ویژگی های داده های اصلی هستند، اما تفسیر آنها به آسانی نیست از دیگر معایب آن می توان به مبادله بین از دست دادن اطلاعات و کاهش ابعاد اشاره کرد پیشنهاد می‌شود از دیگر روش‌های کاهش بعد برای پیش‌پردازش مجموعه داده شامل kernel-pca استفاده کرد. در طراحی شبکه تعداد لایه‌ها افزایش داده شد و از الگوریتم‌های بهینه adam در طراحی آن استفاده کردیم و انواع روش‌های مقداردهی را امتحان ، پیشنهاد می کنیم از دیگر روش‌های یادگیری عمیق استفاده شود .

٦. منابع

[1] H. Rajadurai and U. D. Gandhi, “An empirical model in intrusion detection systems using principal component analysis and deep learning models,” *Comput. Intell.*, vol. 37, no. 3, pp. 1111–1124, 2021, doi: 10.1111/coin.12342.

[2] M. E. Elhamahmy, H. N. Elmahdy, and I. A. Saroit, “A New Approach for Evaluating Intrusion Detection System,” *[1] M. E. Elhamahmy, H. N. Elmahdy, I. A. Saroit, “A New Approach Eval. Intrusion Detect. Syst. vol. 2, no. 11, 2010.*, vol. 2, no. 11, 2010.

[3] W. C. Lin, S. W. Ke, and C. F. Tsai, “CANN: An intrusion detection system based on combining cluster centers and nearest neighbors,” *Knowledge-Based Syst.*, vol. 78, no. 1, pp. 13–21, 2015, doi: 10.1016/j.knosys.2015.01.009.

[4] ”.م. لطیف, “بررسی و مقایسه روش های تشخیص نفوذ در شبکه های کامپیوتری

[5] S. Sunita, B. J. Chandrakanta, and R. Chinmayee, “A Hybrid Approach of Intrusion Detection using ANN and FCM,” *Eur. J. Adv. Eng. Technol.*, vol. 3, no. 2, pp. 6–14, 2016.

[6] Y. Yu and N. Bian, “An Intrusion Detection Method Using Few-Shot Learning,” *IEEE Access*, vol. 8, no. 1, pp. 49730–49740, 2020, doi: 10.1109/ACCESS.2020.2980136.

[7] C. Liu, Z. Gu, and J. Wang, “A Hybrid Intrusion Detection System Based on Scalable K-Means+ Random Forest and Deep Learning,” *IEEE Access*, vol. 9, pp. 75729–75740, 2021, doi: 10.1109/ACCESS.2021.3082147.

[8] S. M. Kasongo and Y. Sun, “A Deep Learning Method With Filter Based Feature Engineering for Wireless Intrusion Detection System,” *IEEE Access*, vol. 7, pp. 38597–38607, 2019, doi: 10.1109/ACCESS.2019.2905633.

[9] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, “Deep Learning Approach for Intelligent Intrusion Detection System,” *IEEE Access*, vol. 7, pp. 41525–41550, 2019, doi: 10.1109/ACCESS.2019.2895334.

[10] B. Singh and A. Kr Ahlawat, “Innovative Empirical Approach for Intrusion Detection Using ANN,” *Int. J. Innov. Res. Comput. Sci. Technol.*, no. 4, pp. 2347–5552, 2016.

[11] C. Xu, J. Shen, X. Du, and F. Zhang, “An Intrusion Detection System Using a Deep Neural Network with Gated Recurrent Units,” *IEEE Access*, vol. 6, pp. 48697–48707, 2018, doi: 10.1109/ACCESS.2018.2867564.

[12] M. Al-Qatf, Y. Lasheng, M. Al-Habib, and K. Al-Sabahi, “Deep Learning Approach Combining Sparse Autoencoder with SVM for Network Intrusion Detection,” *IEEE Access*, vol. 6, pp. 52843–52856, 2018, doi: 10.1109/ACCESS.2018.2869577.

[13] B. Yan and G. Han, “Effective Feature Extraction via Stacked Sparse Autoencoder to Improve Intrusion Detection System,” *IEEE Access*, vol. 6, pp. 41238–41248, 2018, doi: 10.1109/ACCESS.2018.2858277.

[14] م. وزان, *یادگیری عمیق:اصول، مفاهیم و رویکردها* , چاپ اول .no. March ,.تهران: میعاد اندیشه, ۲۰۲۱.

[15] ”.wknnز. استادرحیمي, “تشخیص نفوذ در شبکه بر اساس روش

[16] Y. Al-Nashif, A. A. Kumar, S. Hariri, G. Qu, Y. Luo, and F. Szidarovsky, “Multi-level intrusion detection system (ML-IDS),” *5th Int. Conf. Auton. Comput. ICAC 2008*, pp. 131–140, 2008, doi: 10.1109/ICAC.2008.25.

[17] M. Ahmed, A. Naser Mahmood, and J. Hu, “A survey of network anomaly detection techniques,” *J. Netw. Comput. Appl.*, vol. 60, pp. 19–31, 2016, doi: 10.1016/j.jnca.2015.11.016.

[18] ز. استادرحیمي, “شخیص نفوذ در شبکه بر اساس روش”.Wknn

[19] ه. د. م. ه. مارک, *طراحی شبکه‌های عصبی*. .

[20] S. Rajagopal, P. P. Kundapur, and K. S. Hareesha, “A Stacking Ensemble for Network Intrusion Detection Using Heterogeneous Datasets,” *Secur. Commun. Networks*, vol. 2020, 2020, doi: 10.1155/2020/4586875.

[21] N. Thapa, Z. Liu, D. B. Kc, B. Gokaraju, and K. Roy, “Comparison of machine learning and deep learning models for network intrusion detection systems,” *Futur. Internet*, vol. 12, no. 10, pp. 1–16, 2020, doi: 10.3390/fi12100167.

[22] J. Zhang, F. Li, H. Zhang, R. Li, and Y. Li, “Intrusion detection system using deep learning for in-vehicle security,” *Ad Hoc Networks*, vol. 95, p. 101974, 2019, doi: 10.1016/j.adhoc.2019.101974.

[23] M. A. Ferrag, L. Maglaras, S. Moschoyiannis, and H. Janicke, “Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study,” *J. Inf. Secur. Appl.*, vol. 50, p. 102419, 2020, doi: 10.1016/j.jisa.2019.102419.

[24] S. U. Jan, S. Ahmed, V. Shakhov, and I. Koo, “Toward a Lightweight Intrusion Detection System for the Internet of Things,” *IEEE Access*, vol. 7, pp. 42450–42471, 2019, doi: 10.1109/ACCESS.2019.2907965.

[25] H. Pajouh and A. Choo, “A two-layer dimension reduction and two- tier classification model for anomaly- based intrusion detection in IoT backbone networks Title A two-layer dimension reduction and two-tier classification model for anomaly-based intrusion detection in IoT backbon,” *Ieeexplore.Ieee.Org*, 2016, [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7762123/.

[26] H. Yao, D. Fu, P. Zhang, M. Li, and Y. Liu, “MSML: A novel multilevel semi-supervised machine learning framework for intrusion detection system,” *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1949–1959, 2019, doi: 10.1109/JIOT.2018.2873125.

[27] F. Salo, A. B. Nassif, and A. Essex, “Dimensionality reduction with IG-PCA and ensemble classifier for network intrusion detection,” *Comput. Networks*, vol. 148, no. November, pp. 164–175, 2019, doi: 10.1016/j.comnet.2018.11.010.

[28] S. M. Kasongo and Y. Sun, “A deep learning method with wrapper based feature extraction for wireless intrusion detection system,” *Comput. Secur.*, vol. 92, p. 101752, 2020, doi: 10.1016/j.cose.2020.101752.

[29] A. Yulianto, P. Sukarno, and N. A. Suwastika, “Improving AdaBoost-based Intrusion Detection System (IDS) Performance on CIC IDS 2017 Dataset,” *J. Phys. Conf. Ser.*, vol. 1192, no. 1, 2019, doi: 10.1088/1742-6596/1192/1/012018.

[30] M. Haggag, M. M. Tantawy, and M. M. S. El-Soudani, “Implementing a deep learning model for intrusion detection on apache spark platform,” *IEEE Access*, vol. 8, no. Dl, pp. 163660–163672, 2020, doi: 10.1109/ACCESS.2020.3019931.

[31] D. J. Bartholomew, “Principal components analysis,” *Int. Encycl. Educ.*, pp. 374–377, 2010, doi: 10.1016/B978-0-08-044894-7.01358-0.

[32] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, “A Detailed Analysis of the KDD CUP 99 Data Set,” no. Cisda, pp. 1–6, 2009.

[33] س. پوستفروشان and م. آ. صرام, “به کارگیری الگوریتم بهینه سازی Pso به منظور بهبود طول عمر شبکه های حسگر بیسیم,” *مجله علمی رایانش نرم و فناوری اطلاعات*, vol. 5, no. 3, pp. 55–64, 2016, [Online]. Available: http://jscit.nit.ac.ir/article_51679.html%0Ahttp://jscit.nit.ac.ir/article_51679_90b6a5fc53f91219539365dc1f843f6e.pdf.

[34] Q. A. Al-Haija and S. Zein-Sabatto, “An efficient deep-learning-based detection and classification system for cyber-attacks in iot communication networks,” *Electron.*, vol. 9, no. 12, pp. 1–26, 2020, doi: 10.3390/electronics9122152.

Abstract

One of the important issues with regard to the widespread use of computer networks is cyber attacks and the discussion of security in these networks and databases. Penetration into computer networks with various motives including political, military, financial or showing laxity and weakness in security in There are existing programs that other common techniques become ineffective due to the volume and new malicious features that are increasing exponentially, and traditional methods are not able to maintain security. As a result, intrusion detection systems are used to detect threats and identify attacks from both sides. It has been introduced that it is a tool used to provide communication in information systems, whose main purpose is not to prevent attacks, but to identify attacks and learn their behavior patterns is one of the tasks of this system. One of the most challenging issues in these systems is accuracy, which has attracted the attention of researchers in this field, and with the improvement of accuracy in these systems, the efficiency also increases.
In order to improve the accuracy of intrusion detection systems, several methods have been presented, but there is still a need for research and study in the field of accuracy. In this thesis, a method to solve the problem of accuracy by reducing the dimensions of the NSL-KDD data set with principal component analysis (PCA) which effectively affects the improvement of data processing speed through feature extraction and the use of multi-layer propstern networks, which is one of the learning methods It is deep (DL) by increasing the number of hidden layers from one layer [1] to two hidden layers, which the architecture of this four-layer proceptron (MLP) includes one input layer, two hidden layers and one output layer, causes more efficient learning in these networks and We are trying to improve the accuracy of intrusion detection in this network with the optimal ADAM algorithm and the selection of the LEAKY RELU activator function. The proposed method has been implemented in the Python simulator and compared with the latest works. The proposed method shows that the accuracy for identifying 5 classes in the NSL-KDD dataset has increased to 93%, which shows the greater efficiency of the proposed method compared to the work done**.**

Key words: Intrusion Detection System (IDS), Deep Learning (DL), Multilayer Proceptron (MLP), ADAM Algorithm, Principal Component Analysis (PCA), NSL-KDD Dataset, Accuracy

Salman Higher Education Institute

Faculty of Engineering

Master's thesis in the field of computer engineering, majoring in computer networks

Improving the accuracy of intrusion detection in the network with the help of deep learning with ADAM optimal algorithm and LEAKY RELU activation function

by
Maryam Mahdavi

Supervisor
Mr. Dr. Seyed Reza Kamel Tabakh Farizani

September2022