سامانه‌های پردازشی و ارتباطی چندرسانه‌ای هوشمند
Intelligent Multimedia Processing and Communication Systems (IMPCS)

Fall 2024, 5 (3), 69-86
DOR:

Received: 9 July 2024
Accepted: 20 Aug 2024

The Integrated Three-dimensional Deep Learning Approach for an Efficient Intrusion Detection System Using Spatiol-Temporal Features

Roya Zareh Farkhady[1], Kambiz Majidzadeh[2*], Mohammad Masdari[3], Ali Ghaffari[4]

1. Ph.D. Student, Department of Computer Engineering, Urmia Branch, Islamic Azad University, Urmia, Iran.
2. Assistant Professor, Department of Computer Engineering, Urmia Branch, Islamic Azad University, Urmia, Iran: **Corresponding Author*, kambiz.majidzadeh@iau.ac.ir
3. Assistant Professor, Department of Computer Engineering, Urmia Branch, Islamic Azad University, Urmia, Iran
4. Associate Professor, Department of Computer Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran

Abstract

Introduction: Intrusion detection systems in network traffic increase network security by detecting abnormal inputs and attacks. Deep learning algorithms are used to learn features in network security for large-scale data. On the other hand, if the input data for a model has high dimensions, there will be more neighbors for each data, which leads to higher accuracy in the model. Simultaneously learning spatial and temporal features for each model is challenging too.

Method: This study presents a method for converting input data into three-dimensional. In the proposed model, Convolution with Short-term Long-term Memory Branches (CLBS3) is used to improve features' spatial learning with three-dimensional input data. In parallel, Short-term Long-term Memory learns hierarchical relationships between different features and extracts temporal features. Finally, the integrated approach of the CLBS3 model uses extracted spatial-temporal features for network data classification

Results: Tests were conducted on the UNSW-NB15 dataset, and performance evaluation of CLBS3 shows that compared to contemporary intrusion detection methods, the proposed model with an accuracy of 98.46% with fewer errors of 1.8% in the UNSW-NB15 dataset provides better performance in intrusion detection.

Discussion: The present study aims to propose CLBS method, which detects attacks by considering the limitations of IoT resources. In order to create an efficient and accurate IDS, the combination of convolution neural networks and Short-term Long-term Memory applied in the fog to separate the attacks from normal traffic. Our proposed method was tested using the UNSW-NB15 dataset, and the results demonstrate enhanced accuracy compared to existing methods, as well as a low false positive rate.

Keywords: Deep learning, convolutional networks, three-dimensional, short-term long-term memory, spatial feature, temporal feature.

سامانه‌های پردازشی و ارتباطی چندرسانه‌ای هوشمند
Intelligent Multimedia Processing and Communication Systems (IMPCS)
انجمن علمی تجارت الکترونیکی ایران
واحد ارومیه

رویکرد یادگیری عمیق سه‌بعدی یکپارچه برای سیستم تشخیص نفوذ کارآمد با استفاده از ویژگی‌های فضایی-زمانی

تاریخ دریافت: ۲۹/ ۰٤/ ۱٤۰۳
تاریخ پذیرش: ۳۰/ ۰۵/ ۱٤۰۳

دورهٔ پنجم، پاییز ۱٤۰۳
شمارهٔ، صص: ٦۹-۸٦

رویا زارع فرخادی [۱]، کامبیز مجیدزاده [۲*]، محمد مصدری [۳]، علی غفاری [۴]

۱- دانشجوی دکتری، گروه کامپیوتر، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران، Roya.farkhady@gmail.com
۲- استادیار، گروه کامپیوتر، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران، kambiz.majidzadeh@iau.ac.ir
۳- استادیار، گروه کامپیوتر، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران، Mo.Masdari@iau.ac.ir
۴- دانشیار، گروه کامپیوتر، واحد تبریز، دانشگاه آزاد اسلامی، تبریز، ایران، A.Ghaffari@iaut.ac.ir

چکیده: سیستم‌های تشخیص نفوذ در ترافیک شبکه با تشخیص ورودی‌های ناهنجار و حمله، امنیت شبکه‌ها را افزایش‌می‌دهند. الگوریتم‌های یادگیری عمیق به دلیل توانایی یادگیری ویژگی‌ها در امنیت شبکه برای داده‌های مقیاس وسیع مورداستفاده قرارمی‌گیرند. از طرفی اگر داده‌های ورودی مدلی، دارای ابعاد بالا با شد به ازای هر داده، هم سایگان بی شتری وجود خواهنددا شت که همین امر باعث دقت بی شتر در مدل می شود. ازاین‌رو در این مقاله رو شی برای تبدیل داده‌های ورودی به سه‌بعدی ارائه شده‌ا ست. شبکه‌های ع صبی کانولو شن به‌طور مؤثر ویژگی‌های ف ضایی را به د ست می‌آورند، شبکه‌های ع صبی حافظه کوتاه‌مدت بلند بهتر برای ویژگی‌های زمانی عمل‌می‌کنند. یکپارچه سازی این مدل‌ها می‌تواند قابلیت بهبود سی ستم ت شخیص نفوذ مقیاس و سیع را دا شته‌با شد. یادگیری هم‌زمان ویژگی‌های فضایی و زمانی اطلاعات برای هر مدل، امری چالش‌برانگیز ا ست. در مدل پی شنهادی کانولو شن با حافظه طولانی کوتاه‌مدت دوشاخه با داده‌های سه‌بعدی (CLBS3[۱]) از یک کانولوشن بهبودیافته با ورودی سه‌بعدی برای یادگیری ویژگی‌های فضایی ا ستفاده‌می شود. به صورت هم‌زمان و موازی، حافظه طولانی کوتاه‌مدت، روابط سل سله‌مراتبی بین ویژگی‌های مختلف را یادمی‌گیرد و ویژگی‌های زمانی را استخراج‌می‌کند. درنهایت، رویکرد یکپارچه مدل CLBS3 از ویژگی‌های فضایی-زمانی استخراج‌شده برای دسته‌بندی داده‌های شبکه استفاده‌می‌کند. آزمون‌ها بر روی مجموعه داده UNSWNB15 انجام‌شده است و ارزیابی عملکرد CLBS3 نشان‌می‌دهد که در مقایسه با روش‌های ت شخیص نفوذ معا صر، مدل پی شنهادی با دقت ۹۸/۴۶ با تعداد کمتری ا شتباهات برابر ۰/۱ در مجموعه داده UNSWNB15 و دقت ۹۸/۲۶ با تعداد کمتری اشتباهات برابر ۰/۲ در مجموعه داده TONIOT عملکرد بهتری در تشخیص نفوذ ارا ئه‌می‌د هد. در مجمو عه داده UNSWNB15 مدت‌ز مان اجرای مدل برای پردازش داده ها برابر با ۸۱/. ثانیه اندازه‌گیری شد. م صرف CPU در طی اجرای مدل برابر با تقریباً ۰٪ گزارش شد. این مقدار نشان‌می‌دهد که مدل، در شرایط فعلی اجرا، تقریباً هیچ بار پردازشی قابل‌توجهی بر CPU سیستم واردنمی‌کند.

واژه‌های کلیدی: یادگیری عمیق، شبکه‌های کانولوشنی، سه‌بعدی، حافظه طولانی کوتاه‌مدت، ویژگی فضایی، ویژگی زمانی.

۱. مقدمه

اخیراً، با توجه به تحولات و پیشـــرفت فنّاوری، خدمات آنلاین به‌طور فراوان افزایش‌یافته‌اســـت. انجام خدمات آنلاین به وجود فناوری‌های شبکه وابسته است. اجزای حیاتی یک شبکه ارتباطی شامل محرمانگی، یکپارچگی و در دســـترس بودن هســـتند و هر فعالیت غیرمعمولی که حداقل یکی از این ویژگی‌ها را نقض‌کند، به‌عنوان یک حمله به شـــبکه در نظر گرفته‌می شود؛ بنابراین، برر سی امنیت شبکه امری ضروری و حیاتی است.

سی ستم ت شخیص نفوذ معمولاً با چارچوب امنیتی شبکه ترکیب می شود تا بتواند به ت شخیص حملات نفوذی در شبکه کمک‌کند. این سیستم به‌صورت بازرس عمل‌می‌کند و تمام بسته‌های ترافیک شبکه را برای یافتن نشانه‌های نفوذ برر سی می‌کند. هرچقدر سیستم تشخیص نفوذ قوی‌تر باشـد، می‌تواند اقدامات بیشـتری را برای تشـخیص نفوذ انجام‌دهد و پس از تشخیص، به‌صورت خودکار سرور را از وجود نفوذ در شـبکه آگاه‌سـازد. به‌طورکلی، یک سـیسـتم تشـخیص نفوذ از ترکیب مجموعه‌ای از سیاست‌های امنیتی، نرم‌افزار و دستگاه‌های سخت‌افزاری استفاده می‌کند تا نظارت پیشرفته‌تری را فراهم کند.

سیستم تشخیص نفوذ به دودسته کلی تشخیص مبتنی بر امضا و مبتنی بر رفتار تقسیم‌می‌شود. در سیستم مبتنی بر امضا، درصورتی‌که بسته ترافیک م شکوک به نفوذ با شد، ام ضای آن با پایگاه داده ام ضای حملات مقایسه شده و در صورت تطبیق، سیستم هشدارمی‌دهد. این نوع سیستم به‌عنوان تشخیص مبتنی بر دانش نیز شناخته‌می‌شود، زیرا از پایگاه داده برای تشـــخیص اســـتفاده‌می‌کند. این روش دقت بالا در تشخیص با حداقل خطای مثبت دارد ولی در تشخیص حملات جدید و نامعلوم ممکن اسـت دچار مشـکل شـود. روش مبتنی بر رفتار بسـته ترافیک را با الگوهای نرمال مقایسه‌می‌کند، اما اگر نمونه‌ای حالت نرمال را تقلیدکند، قادر به تفکیک حمله وارد شده از حالت نرمال نی ست. این روش در تشـــخیص حملات جد ید از روش مبتنی بر امضـــا بهتر عمل‌می‌کند. هر یک از این رویکردها قابلیت‌ها و محدودیت‌های خود رادارند، اما ترکیب آن‌ها می‌تواند به‌دقت و کارایی بالاتر در تشـــخیص حملات کمـک‌کنـد. مـدل تشـــخیص نفوذ کـه در این مقـالـه پی شنهاد شده‌ا ست، بر ا ساس ترکیب هر دو رویکرد ت شخیص ام ضا و رفتار است.

رویکردهای سنتی مانند رمزنگاری، دیواره‌های آتش و نرم‌افزارهای مقابله با حملات، ازجمله روش‌های تشـــخیص حملات بودند؛ اما این روش‌ها در داده‌های با ابعاد کم عملکرد خوبی داشتند، ولی در داده‌های بزرگ‌تر و حملات مکرر مانند حملات انکار سـرویس که هکرها پشـت سـر هم حمله‌می‌کنند، قادر به مقابله با این مشـکلات نیسـتند. بیشـتر تحقیقـات اخیراً از یـادگیری مـاشـــین برای تشـــخیص حملات استفاده‌کرده‌اند.

بااین‌وجود، الگوریتم‌های معمول یادگیری ماشـــین محدودیت‌های جدی در یادگیری ویژگی‌های یکپارچه حملات دارند[۱]. به‌علاوه، این الگوریتم‌ها عملکرد کمتری برای داده‌های با ترافیک نویزی و ابعاد بالا ارائه‌می‌دهند. الگوریتم‌های یادگیری ماشـــین که از ترکیب چند روش یادگیری ماشین برای سیستم تشخیص نفوذ استفاده‌می‌کنند، همچنین با مشـــکل پیچیدگی بالا روبه‌رو هســـتند و با داده‌های ابعاد بالا نیز مشکل‌دارند که باعث کاهش کارایی آن‌ها می‌شود.

اخیراً الگوریتم‌های یادگیری عمیق در ت شخیص نفوذ پی شرفت‌های قابل‌توجهی داشـته‌اند. مدل‌های تشـخیص نفوذ با اسـتفاده از یادگیری عمیق که دارای ساختار غیرخطی هستند، عملکردهای یادگیری بهتری ارائه‌کرده و دقت ت شخیص نفوذ را افزایش‌داده‌اند. از مدل‌های غیرخطی رایج می‌توان به شـبکه‌های عصـبی بازگشـتی[۲]، کانولوشـنی[۳]، و حافظه طولانی کوتاه‌مدت [۴] اشاره‌کرد.

مدل‌های مبتنی بر شبکه‌های کانولوشنی از یک معماری مشترک و بی‌تغییر جابه‌جایی بهره‌می‌برند تا لایه‌های پنهان و ویژگی‌های دیگر را ا ستخراج‌کنند. آن‌ها از شبکه‌های چندلایه‌ای پر سپترون و شبکه‌های کاملاً متصل برای تعیین کلاس حمله استفاده‌می‌کنند. بااین‌حال، برخی از این شبکه‌ها به بیش برازش منجرمی‌شوند.

مدل‌های ت شخیص نفوذ بر پایه شبکه‌های ع صبی بازگ شتی برای پردازش دنباله‌های طولانی‌مدت داده‌های ترافیک ورودی، از ســـاختار شبکه عصبی پیشرو با حافظه داخلی استفاده‌می‌کنند. این مدل‌ها برای کنترل وضـــعیت عملکرد و بهبود نرخ یادگیری از حلقه‌های بازخورد اســـتفاده‌می‌کنند. حافظه طولانی کوتاه‌مدت یک مدل شـبکه عصـبی بازگ شتی ا ست که می‌تواند برای ت شخیص نفوذ در دنباله‌های داده‌های ترافیک استفاده‌شود. این مدل شامل اتصالات بازخوردی است که برای شـناسـایی نفوذ با ویژگی‌های زمانی مناسـب هسـتند. سـاختار عمیق غیرخطی شبکه‌های ع صبی کانولو شنی و حافظه طولانی کوتاه‌مدت به استخراج ویژگی‌های ترافیک‌های شبکه کمک‌می‌کند. شبکه‌های عصبی کـانولوشـــنی بـا یـادگیری ویژگی‌هـای فضـــایی عمیق از طریق همبستگی‌های فضایی و با نرخ یادگیری سریع، دقت شنا سایی بالایی ارائه می‌دهند؛ اما اثبات‌شـده اسـت که شـبکه‌های عصـبی کانولوشـنی به‌طور خودکار ویژگی‌های زمانی را یادنمی‌گیرند و نیاز به یک سـاختار اضـافی برای یادگیری وابسـتگی‌های طولانی‌مدت در ویژگی‌های زمانی دارند. حافظه طولانی کوتاه‌مدت یک کلاس ویژه از شـبکه‌های عصـبی بازگشتی است که می‌تواند ویژگی‌های زمانی از ترافیک‌های شبکه را از طریق حـافظـه طولانی کوتـاه‌مـدت خود بـه‌طور مؤثرتر یـادبگیرد. بااین‌وجود، شـبکه حافظه طولانی کوتاه‌مدت نیازمند زمان بیشـتر برای آموزش است و نسبت به شبکه‌های عصبی کانولوشنی به منابع بیشتری نیاز دارد و دانش محدودی از اطلاعات گذشته دارد[۵، ۶].

تحقیقات نشـان می‌دهد که شـبکه‌های عصـبی کانولوشـنی دارای کارایی بالا و زمان آموزش سـریع اسـت، درحالی‌که حافظه طولانی کوتاه‌مدت دارای پردازش دقیق و مؤثر ویژگی‌های زمانی اسـت. ادغام این دو مدل می‌تواند به بهترین نحو از مزایا هرکدام استفاده‌کند و حتی بر غلبه بر محدودیت‌های آن‌ها کمک‌کند.

بر همین اســاس یک مدل یادگیری عمیق ترکیبی در این مقاله پیشنهاد شده است که در آن شبکه‌های کانولوشنی و حافظه طولانی کوتاه‌مدت به صورت موازی برای استخراج ویژگی‌های فضایی – زمانی مجموعه داده‌ها مدل می‌شوند.

بیشــتر مقالات شــبکه‌های عصــبی عمیق را به‌صــورت یک‌بعدی طراحی و پیاده‌ســازی کرده‌اند، این در حالی اســت که تبدیل داده‌ها به ســه‌بعدی برتری قابل‌توجهی نســبت به حالت یک‌بعدی دارد که این مفهوم و برتری در مدل پیشــنهادی ما ایجادشــده‌اســت. شــبکه‌های کانولوشنی سه‌بعدی دارای دقت بیشتری نسبت به بُرداری و دوبعدی هستند. چون در این شبکه‌ها امکان استفاده از جزئیات بیشتری در تصمیم‌گیری وجوددارد. با توجه به اینکه در مسئله تشخیص نفوذ در شبکه، دقت تشخیص پارامتر مهمی است، بنابراین از تبدیل سه‌بعدی ورودی در جهت افزایش دقت استفاده‌شده‌است.

در روش پیشنهادی، ابتدا داده‌های یک‌بعدی با استفاده از تکنیک جایگشت به داده‌های دوبُعدی تبدیل‌می‌شوند و سپس با استفاده از لایه‌های شــبکه کانولوشــنی دوبعدی با اندازه هســته یک، داده‌ها به داده‌های سه‌بعدی تبدیل‌می‌شوند. در گام بعدی، از جداکننده کانال‌ها و تکنیک جایگشت برای ایجاد دوشاخه متفاوت استفاده‌می‌شود و سپس داده‌های سه‌بعدی توسط دوشاخه متفاوت لایه‌های شبکه کانولوشنی و لایه‌های شبکه عصبی بازگشتی استخراج ویژگی می‌شوند. در این مقاله، موارد زیر ایجادشده‌است:

- تبدیل داده‌های ورودی به داده‌های سه‌بُعدی
- اســتفاده از تکنیک‌های جداکننده کانال‌ها و تکنیک جایگشــت برای ایجاد دوشاخه متفاوت
- ترکیب روش شبکه عصبی کانولوشن و حافظه طولانی کوتاه‌مدت در دو شاخه موازی
- استخراج ویژگی‌های فضایی و زمانی
- افزایش دقت تشخیص و کاهش نرخ مثبت کاذب
- مقایسه مدل پیشنهادی با سیستم‌های جدید تشخیص نفوذ در شبکه‌های اینترنت اشیا

مقاله به صورت مرور کارهای پیشین در بخش دوم و توضیح روش پیشنهادی و پیاده سازی‌های آن در بخش سوم بیان شده است. نتایج عملکرد و بررسی مقایسه‌ای در بخش چهارم تشریح شده‌اند. در بخش پنجم تحلیل مدیریت و نرخ خطا و در بخش ششم تحلیل مصرف منابع پرداز شی و حافظه روش پیشنهادی بررسی شده‌است، در بخش هفتم نتیجه‌گیری و پیشنهادها برای آینده ذکرشده‌اند.

۲. کارهای پیشین

تحقیقات بسیار زیادی برای تشخیص نفوذ در شبکه‌ها با استفاده از روش‌های یادگیری ماشین و یادگیری عمیق انجام‌شده‌است. در این بخش، بررسی شده‌است که چگونه این روش‌ها می‌توانند بهبود بخشی در تشخیص نفوذ داشته‌باشند. سه دسته اصلی شامل روش‌های مبتنی بر یادگیری ماشــین، روش‌های مبتنی بر ترکیب روش‌های یادگیری ماشین و روش‌های مبتنی بر یادگیری عمیق بررسی شده‌است. البته فقط به بررسی چند نمونه اخیر پرداخته‌شده‌است.

۲,۱. روش‌های مبتنی بر یادگیری ماشین

در دهه گذشته، استفاده از الگوریتم‌های یادگیری ماشین و یادگیری عمیق برای مدل‌های تشــخیص نفوذ به‌طور قابل‌توجهی افزایش‌یافته اســت. توانایی ارائه پیش‌بینی دقیق با اســتفاده از یادگیری ویژگی‌های مؤثر، از مزایای اصــلی این مدل‌های دســتگاه‌های تشــخیص نفوذ بوده‌است. بیشتر مدل‌های سیستم‌های تشخیص نفوذ موجود بر اساس مدل‌های یادگیری نظارت شده ساخته شده‌اند. در این قسمت به برخی از این تحقیقات اشاره شده‌است. ماشین بردار پشتیبان به‌طور گسترده برای این دســتگاه‌ها اســتفاده‌می‌شــود. در مطالعه [۷] نویســندگان از ماشــین بردار پشــتیبان با افزایش ویژگی برای تشــخیص نفوذ اســتفاده‌کردند و آن را بر روی مجموعه داده NSL-KDD با دقت ۹۹/۱۸٪، نرخ تشخیص ۹۹/۸۵٪ و نرخ اشتباه ۲/۹۶٪ ارزیابی‌کردند. همچنین، در مطالعه [۸] از مدل مبتنی بر ماشین بردار پشتیبان برای تشخیص نفوذ استفاده‌شد که بر روی مجموعه داده AWID بررسی‌شد و دقت ۹۹/۲۵٪ و نرخ خطای مثبت کاذب ۱/۲٪ داشت.

در مطالعه [۹] از مدل k نزدیکترین همســایه برای ســاخت فیلتر هشدار مدل سیستم تشخیص نفوذ استفاده شد که بر روی مجموعه داده DARPA1999 با دقت ۸۸/۶٪ تســت شــد. همچنین در مطالعه [۱۰] نیز از k نزدیک‌ترین همسایه برای مدل سیستم تشخیص نفوذ از طریق انتخاب ویژگی تحلیل اصلی مؤلفه برای مجموعه داده ADFA-LD استفاده شد و دقت ۹۹/۵٪، نرخ واقعی مثبت ۹۹/۸٪، نرخ واقعی منفی ۹۹/۷٪، نرخ خطای مثبت کاذب ۰/۳٪ و نرخ خطای منفی کاذب ۰/۲٪ را به دست آورد.

در مطالعه[۱۱] الگوریتم K-means برای تشخیص نفوذ از مجموعه داده KDD99 اعمال شد و دقت تشخیص ۹۹/۸٪ و نرخ خطای مثبت کاذب ۰/۳٪ به‌دست آمد. سیستم‌های تشخیص نفوذ در مطالعه [۱۲] با استفاده از جنگل‌های تصادفی پیاده‌سازی شدند و به‌دقت ۹۹/۶۷٪ و نرخ خطای مثبت کاذب ۰/۸٪ را بر روی مجموعه داده NSL-KDD به دست آورند. در مطالعه [۱۳] از درخت تصمیم برای تشخیص نفوذ اســتفاده‌شــد و در مجموعه داده KDD99 نفوذها با دقت ۸۸٪ شناسایی‌شدند. در مطالعه [۱۴]از رگرسیون لجستیک بر روی مجموعه داده NSL-KDD استفاده‌شد و حملات با دقت ۹۷/۵٪ شناسایی‌شدند.

بااینکه الگوریتم‌های یادگیری ماشــین می‌توانند نرخ دقت کلی بالایی را به دست آورند، اما برای بسیاری از انواع حملات، به‌خصوص U2R و R2L نرخ تشــخیص پایینی دارند. همچنین تعادل بین مثبت و منفی‌های کاذب در بیشتر روش‌ها نامتقارن است، به‌طوری‌که مثبت‌های کاذب با کاهش منفی‌های کاذب افزایش می‌یابند.

۲،۲. روش‌های مبتنی بر ترکیب روش‌های یادگیری ماشین

بعضی از تحقیقات از الگوریتم‌های یادگیری ماشین ترکیبی و پیشرفته بهره گرفتند تا میزان مثبت‌های کاذب را کاهش دهند. در مطالعه [۱۵]، یک سیستم تشخیص نفوذ با استفاده از ترکیب ماشین بردار پشتیبان و k نزدیک‌ترین همسایه طراحی‌شد که با دقت ۸۹/۰۲٪ و نرخ خطا ۱۲٪ بر روی مجموعه داده KDD99 مورد ارزیابی قرارگرفت. در مطالعه [۱۶] ماشین بردار پشتیبان و ژنتیک با یکدیگر ترکیب و برای تشخیص نفوذ استفاده‌شدند. این مدل بر روی مجموعه داده KDD99 توسعه‌یافت و با ۹۶/۶۱٪دقت، ۰/۰۷٪ منفی‌های کاذب و ۳/۳۹٪ مثبت‌های کاذب و ۲/۴٪نرخ خطا به دست آمد. در مطالعه[۱۷] یک سیستم تشخیص نفوذ مبتنی بر ماشین بردار پشتیبان و C5 پیشنهادشد که دقت ۹۴٪ به‌دست آمد.

تکنیک‌های تشخیص نفوذ مبتنی بر یادگیری ماشین پیشرفته شامل تکنیک‌های مبتنی بر شبکه‌های عصبی است. محققان در مطالعه [۱۸] از شبکه‌های عصبی مصنوعی بر مجموعه داده ترافیک شبکه معتبر استفاده کردند و دقت ۹۸٪، حساسیت ۹۵٪ را به دست آوردند. در مطالعه [۱۹] از شبکه‌های عصبی مصنوعی چندکلاسه متوالی روی KDD99 با دقت ۹۸/۲۵٪، نرخ مثبت های کاذب ۳/۷۷٪ و نرخ منفی‌های کاذب ۱/۲۶٪ و روی مجموعه داده UNSWNB15 با دقت ۸۶/۴٪ و نرخ مثبت‌های کاذب ۲/۸ درصد استفاده‌شد. محققان از مدل یکپارچه‌شده از شبکه‌های عصبی با انتخاب ویژگی مبتنی بر همبستگی بر روی مجموعه داده‌های NSL-KDD و UNSWNB15 برای تشخیص نفوذ در [۲۰]استفاده‌کردند. این شبکه عصبی با دقت ۹۸/۴۵٪بر روی مجموعه داده NSL-KDD و دقت ۹۶/۴٪ برای UNSWNB15 به دست آمد.

باوجوداینکه الگوریتم‌های یادگیری ماشین ترکیبی و پیشرفته بهینه‌تر از الگوریتم‌های ماشین معمولی عمل‌می‌کنند، اما محدودیت هایی در انتقال مجموعه داده های بزرگ وجوددارد. این محدودیت‌ها به دلیل معماری کم‌عمق الگوریتم‌هاست که توانایی آن‌ها در استخراج ویژگی‌های عمیق از داده‌های شبکه را کاهش می‌دهد.

۲،۳. روش‌های مبتنی بر یادگیری عمیق

تعداد زیادی تحقیق با استفاده از الگوریتم‌های یادگیری عمیق برای تشخیص نفوذ انجام شده‌است. همان‌طور که قبلاً بیان شد، استخراج ویژگی‌های فضایی-زمانی می‌تواند به طرز قابل‌توجهی باعث بهبود در تشخیص نفوذ شود، اما این چالش‌های خود را دارد. به همین دلیل، در این بخش به بررسی تحقیقاتی که از شبکه‌های کانولوشنی و حافظه طولانی کوتاه‌مدت استفاده‌کرده‌اند، پرداخته‌شده‌است.

در چند سال اخیر، درزمینه تشخیص نفوذ در شبکه‌های اینترنت اشیا، مدل‌های یادگیری عمیق مختلفی معرفی‌شده‌اند. اکثر کارهای اخیر بر روی دسته‌بندی دودویی بر مجموعه داده‌های مختلف با نتایج موفق تمرکز دارند. بااین حال، اکثر این تحقیقات نمی‌توانند هم دقت بیشتری داشته‌باشند و هم امتیاز F1 بالا را حفظ‌کنند و هم منفی کاذب را کاهش دهند.

تعداد زیادی تحقیق با استفاده از الگوریتم‌های یادگیری عمیق برای تشخیص نفوذ انجام‌شده‌است. در مطالعه [۲۱] از یک سیستم تشخیص نفوذ مبتنی بر شبکه‌های عصبی کانولوشن برای تشخیص حملات انکار سرویس استفاده‌شد که با دقت ۹۹٪ بر روی مجموعه داده KDD99 و با دقت ۹۱/۵٪ بر روی مجموعه داده CSE-CIC-IDS2018 عملکرد قابل-قبولی داشت. در مطالعه [۲۲]یک سیستم تشخیص نفوذ مبتنی بر شبکه‌های عصبی کانولوشن و ژنتیک برای مجموعه داده NSL-KDD ارائه شد که حملات با دقت ۹۸/۲٪ تشخیص داده‌شدند. در مطالعه [۲۳] از یک شبکه عصبی حافظه طولانی کوتاه‌مدت برای سیستم تشخیص نفوذ استفاده‌شد که بر روی مجموعه داده CIDDS با دقت ۸۵٪ آزمایش-شد. در مطالعه [۲۴] از یک شبکه عصبی حافظه طولانی کوتاه‌مدت دوجهته برای مجموعه داده ADFA با دقت ۹۰٪ استفاده‌کردند. در مطالعه [۲۵]یک شبکه عصبی حافظه طولانی کوتاه‌مدت بر روی مجموعه داده ISCX-UNB با دقت ۹۷٪ استفاده‌کردند. علاوه بر شبکه‌های عصبی کانولوشنی و شبکه‌های عصبی بازگشتی/حافظه طولانی کوتاه‌مدت، بسیاری از الگوریتم‌های یادگیری عمیق دیگر نیز به‌صورت جداگانه یا به همراه الگوریتم‌های دیگر برای تشخیص نفوذ استفاده‌شده‌اند. بااین‌حال، شبکه‌های عصبی کانولوشنی و حافظه طولانی کوتاه‌مدت همیشه در بیشتر مجموعه داده‌های نفوذ، بهترین تشخیص حمله را ارائه کرده‌اند. آن‌ها همچنین از محدودیت‌های خاصی که در بخش قبل ذکرشده‌اند رنج‌می‌برند. برای کاهش محدودیت‌های آن‌ها، برخی نویسندگان تلاش-کرده‌اند با استفاده از مدل‌های ترکیبی یا مدل‌های یکپارچه اقدام-کنند[۲۶]. نویسندگان در مطالعه [۲۷] شبکه عصبی کانولوشنی حافظه طولانی کوتاه‌مدت برای مدل سیستم تشخیص نفوذ توسعه‌دادند که بر روی مجموعه داده ISCX-UNB با دقت ۹۷/۲۹٪ آزمایش‌شد. همان‌طور که قبلاً توضیح داده‌شد، استخراج ویژگی‌های فضایی-زمانی می‌تواند به طرز قابل‌توجهی تشخیص نفوذ را بهبودبخشد اما این چالش زیادی دارد. نویسندگان در مطالعه [۲۸] از شبکه عصبی عمیق برای یادگیری ویژگی‌های فضایی-زمانی استفاده‌کردند. این روش بر روی مجموعه داده DARPA1998 و ISCX2012 با دقت ۹۹/۸٪ و ۶/٪ ۹۹ اعمال‌شد. به همین ترتیب، در مطالعه [۲۹] مدل ترکیبی از شبکه‌های عصبی کانولوشن برای تشخیص نفوذ توسعه داده‌شد. این مدل بر روی مجموعه داده UNSWNB15 تست شد و دقت ۹۵/۶٪ به دست آورد. بااین‌حال، هر دو این مدل‌ها دچار محدودیت‌هایی در مدیریت مجموعه داده‌های نامتعادل هستند. با در نظر گرفتن محدودیت‌های روش‌های پیشین در تعادل بین دقت و امتیاز f1 و کاهش نرخ مثبت کاذب مدلCLBS3 در این مقاله پیشنهادشد. همچنین در این مدل با بهره‌گیری از مزایای هر دو مدل شبکه‌های عصبی کانولوشنی و حافظه طولانی کوتاه‌مدت برای تشخیص نفوذ سیستم قادر به تشخیص امضا و رفتار است. این کار به‌ویژه

در محیط‌هایی که به دلیل اجرای هم‌زمان برنامه‌های چندگانه، الگوهای ترافیک متعددی تجربه‌می‌کند، مفید خواهدبود.

با توجه به اینکه روش پیشنهادی بر اساس تشخیص NIDS در IoT است، در جدول ۱ تحلیل مقایسه‌ای از کارهای قبلی که از روش یادگیری عمیق برای تشخیص نفوذ در اینترنت اشیا در لایه شبکه استفاده‌کرده‌اند، انجام شده است. در تمامی روش‌های موردبرر سی، از یادگیری عمیق نظارت شده استفاده شده است. جدول بر اساس زمان انتشار روش‌ها تنظیم‌شده‌است.

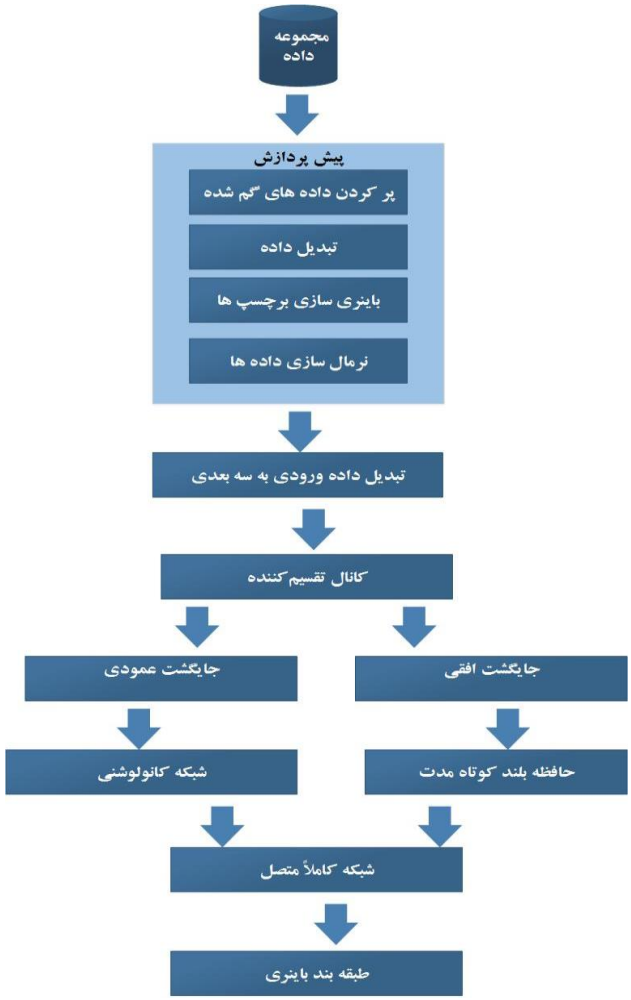
جدول ۱: مقایسه روش‌های تشخیص نفوذ در اینترنت اشیا بر اساس یادگیری عمیق

ردیف	مرجع	سال	مدل	دقت	مجموعه داده
۱	[۳۰]	۲۰۱۶	ANN	۹۹	شبیه‌سازی‌شده
۲	[۳۱]	۲۰۱۷	GRNN	۹۸/۹۱	KDDCUP99
۳	[۳۲]	۲۰۱۸	ANN	۸۹	UNSWNB-15
۴	[۳۳]	۲۰۱۸	DT+NB+ANN	۹۸/۹۷	UNSWNB15
۵	[۳۴]	۲۰۱۸	LSTM	۹۵	UNSWNB-15
۶	[۳۵]	۲۰۱۹	ANN	۹۹/۸۶	KDD CUP99
۷	[۳۶]	۲۰۱۹	ANN	۱۰۰	N_BaIoT
۸	[۳۷]	۲۰۲۰	BILSTM	۹۹/۵۷ ۸۵/۵۹	NSL KDD UNSWNB15
۹	[۳۸]	۲۰۲۰	DT+NB+ANN	۹۹	Modbus network traffic
۱۰	[۳۹]	۲۰۲۰	LCNN + GRNN	۹۱	NSL-KDD
۱۱	[۴۰]	۲۰۲۰	CNN + LSTM	۹۴/۸۰	N_BaIoT
۱۲	[۴۱]	۲۰۲۱	bFNN + mFNN	۹۹/۹۶	BOTIOT
۱۳	[۴۲]	۲۰۲۳	CNN+GRU	۹۹/۹۲	Kddcup99
۱۴	[۴۲]	۲۰۲۳	CNN+LSTM	۹۹/۹۵	Kddcup99
۱۵	[۴۳]	۲۰۲۳	ANN-ICA	۹۰	Kddcup99
۱۶	[۴۴]	۲۰۲۱	FUZZY	۹۱/۲۶	Kddcup99

۳. روش پیشنهادی

در این بخش، هدف طراحی یک سیستم تشخیص نفوذ در شبکه‌های اینترنت اشیا، با استفاده از روش‌های یادگیری عمیق به‌صورت سه‌بعدی است. در مدل پیشنهادی، از تکنیک‌های مختلفی استفاده شده تا دقت و قدرت مدل افزایش یابد. روش جدید یادگیری عمیق سه‌بعدی مبتنی بر تقسیم‌کننده شاخه برای تشخیص نفوذ در شبکه‌های اینترنت اشیا

این مقاله بر تمرکز دسته‌بندی دودویی با استفاده از شبکه‌های کانولوشنی و حافظه طولانی کوتاه‌مدت با تنظیم ورودی سه‌بعدی به نام CLBS3 معرفی‌شده‌است. در شکل ۱، مراحل مختلف مدل پیشنهادی به‌طورکلی نشان داده‌شده‌است که از سه بخش اصلی تشکیل‌شده‌است. در بخش اول، مجموعه داده UNSW-NB 15 و TONIOT برای تشخیص نفوذ استفاده‌شده است. در بخش دوم، از نکنیک‌های پیش‌پردازش برای آماده‌سازی داده‌ها استفاده‌شده‌است. بخش سوم به چارچوب اصلی مدل پیشنهادی به همراه تکنیک‌های جداکننده کانال، نگاشت و لایه‌های مختلف یادگیری عمیق برای استخراج ویژگی می‌پردازد. در ادامه، هر یک از مراحل مدل پیشنهادی در زیر بخش‌های جداگانه توضیح داده‌شده و جزئیات کامل در هر بخش آمده‌است.

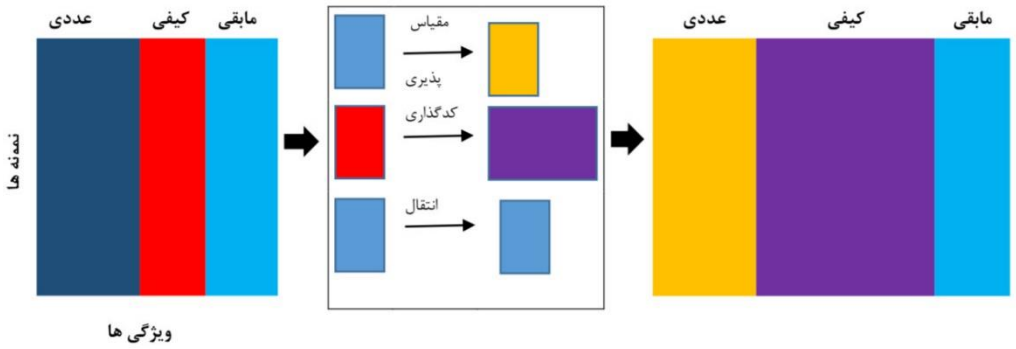


شکل ۱: فلوچارت تشخیص نفوذ مدل CLBS3

۳,۱. پیش‌پردازش

داده‌های اولیه ممکن است حاوی اطلاعات نامعلوم و اغتشاش بوده که بی‌تردید بر کارایی مدل تأثیرگذار خواهدبود. مدل ارائه‌شده نیازمند به‌کارگیری چندین مرحله پیش‌پردازش پیش از استفاده از فن‌های یادگیری عمیق است. این فرآیندهای پیش‌پردازش به بهبود کیفیت

داده‌های مربوط به اینترنت اشیا، ارتقای دقت مدل ارائه‌شده و کاهش پیچیدگی محاسبات کمک‌می‌کنند. مراحل آماده‌سازی داده شامل دو مرحله کلیدی پاک‌سازی و تبدیل داده‌ها است. پس از اجرای این مراحل، همه داده‌ها آماده برای به‌کارگیری در مدل پیشنهادی و دیگر روش‌های مقایسه‌ای می‌شوند. در ادامه این مراحل توضیح داده‌شده‌است.



شکل ۲: تبدیل داده‌ها در مرحله آماده‌سازی

۳،۱،۱ پاک‌سازی داده‌ها

گام نخست در فرایند پیش‌پردازش داده‌ها معمولاً به پاک‌سازی داده‌ها اختصاص دارد. این مرحله شامل فرآیندهایی نظیر دور ریختن داده‌های بی‌ربط، حذف داده‌های تکراری، تکمیل اطلاعات ناقص و مواردی ازاین‌دست است.

- **حذف داده‌های بی‌ربط**

منظور از داده‌های بی‌ربط داده‌هایی هستند که وجود آن ویژگی تأثیری در یادگیری مدل نخواهدداشت و ازنظر فراوانی مقادیر آن ویژگی می‌توان آن را تشخیص‌داد. یکی از روش‌های رایج برای حذف داده‌های بی‌ربط استفاده از خصوصیت منحصربه‌فرد بودن است بدین‌صورت که اگر مقادیر یک ویژگی فقط یک مقدار باشد و به عبارتی واریانس آن خیلی کم باشد، آن ویژگی از مجموعه داده حذف‌خواهدشد. اگر واریانس یک ویژگی برابر تعداد رکوردهای یک مجموعه داده باشد نیز آن ویژگی از مجموعه ویژگی‌ها حذف‌می‌شود مانند ویژگی id در مجموعه داده UNSWNB15 که شماره ردیف را نگهداری‌می‌کند.

- **داده‌های مفقودشده**

برخی از ویژگی‌های بعضی از نمونه‌های مجموعه داده ممکن است دارای مقادیر خالی باشد. این امر به دلیل عوامل زیادی مثل ایجاد نویز در سیستم یا از قلم افتادن در درج مقادیر ویژگی‌ای اتفاق بیافتند. مدیریت و رفع مشکل داده‌های مفقودشده حائز اهمیت است، چراکه مدل‌های پیشنهادی و الگوریتم‌های دیگر نمی‌توانند باوجود داده‌های گمشده به‌درستی آموزش ببینند.

در این مقاله، برای جلوگیری از تأثیر منفی مقادیر گمشده بر مدل، تمامی مقادیر خالی در دیتاست با مقدار مناسبی جایگزین می‌شوند. به این صورت که اگر یک ستون شامل داده‌های متنی (رشته‌ای) باشد، از مقدار مد (mode) آن ستون استفاده‌می‌کنیم و اگر داده‌ها عددی باشند، از میانگین (mean) آن ستون استفاده‌می‌شود.

۳،۱،۲ تبدیل داده‌ها

تبدیل داده‌های بی‌معنا و تبدیل داده‌ها به قالب مناسب برای تجزیه‌وتحلیل است. این مرحله شامل نرمال‌سازی داده‌های عددی در مقیاس مشترک، رمزگذاری متغیرهای طبقه‌بندی یا اعمال تبدیل‌های ریاضی برای دستیابی به خواص توزیع بهتر است. با تبدیل داده‌ها اطمینان داده‌می‌شود که ویژگی‌های مختلف در مقیاس قابل‌مقایسه هستند و سوگیری‌هایی را که ممکن است بر تجزیه وتحلیل تأثیر بگذارند حذف‌می‌شود. تأثیر مراحل تبدیل داده‌ها بر روی تعداد ویژگی‌ها در شکل ۲ نشان داده‌شده‌است.

- **کدگذاری**

همه ورودی‌های استفاده‌شده در متدهای یادگیری ماشین و مدل موردنظر باید در قالب آرایه‌هایی از اعداد صحیح یا اعداد ممیزی شناور باشند. البته، در داده‌های موجود ممکن است ویژگی‌هایی با نوع دسته‌بندی‌شده وجود داشته‌باشند که نیاز به تبدیل به نوع دیگری در این مرحله‌دارند. دو روش متداول برای این تبدیل، استفاده از روش‌های کدگذاری برچسب و کدگذاری وان هات می‌باشد. استفاده از روش کدگذاری وان هات می‌تواند به دلیل افزایش بیش‌ازحد تعداد ویژگی‌ها به نسبت تعداد دسته‌ها، کمتر کارآمد باشد. در این تحقیق روش‌های مختلف کدگذاری بررسی‌شد که در قسمت ارزیابی نمایش داده‌شده است و در بین آن‌ها روش کدگذاری برچسب موجود در کتابخانه sklearn برای تغییر ویژگی‌های با نوع رشته‌ای یا دسته‌بندی‌شده بهترین دقت را در مجموعه داده‌های موردبررسی داشت.

- **نرمال‌سازی**

برای دستیابی به نتایج دقیق‌تر، لازم است که بازه مربوط به ویژگی‌های مختلف، به نحوی با یکدیگر یکسان و یا نزدیک شوند. برای این منظور از روش‌های نرمال‌سازی استفاده‌می‌شود. روش‌های مختلفی برای نرمال‌سازی وجوددارد که در مقالات و تحقیقات مورداستفاده قرارمی‌گیرد. تکنیک‌های مختلفی برای نرمال‌سازی داده‌ها وجوددارد مانند min-max, z-score, tanh, ….

نرمال‌سازی یک فرایند است که در آن، داده‌های مربوط به هر ویژگی به‌گونه‌ای تغییر مقیاس یا تبدیل‌می‌شوند که در یک بازه خاص قرار بگیرند. در این مقاله تمامی این روش‌ها برای نرمال‌سازی اعمال‌شد و تأثیر هریک بر دقت مدل ارزیابی‌شد که در قسمت ارزیابی نتایج به‌دست‌آمده گزارش‌شده‌است. طبق نتایج به‌دست‌آمده روش min-max بهترین دقت را داشته‌است و برای روش پیشنهادی از این روش نرمال‌سازی استفاده‌شده‌است.

روش حداقل-حداکثر یکی از روش‌های معمول برای تغییر داده‌ها به یک مقیاس و بازه نوین است که با استفاده از فرمول ۱ محاسبه می‌شود:

$$X_{norm} = \frac{X_i - Min\ (X_i)}{Max(X_i) - Min\ (X_i)} \qquad (۱)$$

X_i نشان‌دهنده ارزش ویژگی عددی برای نمونه i است، Min و Max نشان‌دهنده حداقل و حداکثر ارزش‌های ویژگی‌های عددی می‌باشند. با استفاده از مقیاس‌بندی، تمامی ویژگی‌ها در دو مجموعه داده ادیری در بازه صفرتا یک تبدیل‌می‌شوند.

۳,۲. مدل CLBS3

روش‌های یادگیری عمیق به‌صورت گسترده در زمینه‌های مختلف ازجمله سیستم‌های تشخیص نفوذ به کار گرفته‌شدند. مزیت اصلی روش‌های یادگیری عمیق استخراج ویژگی‌های مختلف از داده‌های چندبعدی است. شبکه‌های عصبی کانولوشن و حافظه طولانی کوتاه‌مدت دو روش مهم در یادگیری عمیق هستند.

با تحلیل روش‌های قبلی، مشخص‌شد که در اکثر روش‌ها تعادل بین دقت و امتیاز F1 وجودندارد. برای رسیدن به این تعادل و کاهش نرخ مثبت کاذب، مدل CLBS3 در این مقاله پیشنهادشده‌است. این مدل با استفاده از ویژگی‌های شبکه‌های عصبی کانولوشنی و حافظه طولانی کوتاه‌مدت برای تشخیص نفوذ، قادر به تشخیص امضا و رفتار است. این کار به‌خصوص در محیط‌هایی که الگوهای ترافیک متعددی وجود دارد، مفید است. مقاله بر تمرکز دسته‌بندی دودویی با استفاده از شبکه‌های کانولوشنی و حافظه طولانی کوتاه‌مدت با تنظیم ورودی سه‌بعدی تمرکز دارد. درنهایت، یک مدل ترکیبی برای تشخیص حمله با دقت و امتیاز F1 قابل‌توجه ارائه‌می‌دهد و نرخ مثبت کاذب را به‌صورت چشمگیری کاهش می‌دهد.

طبق معماری مدل پیشنهادی **CLBS3** در شکل ۳، داده‌های دوبعدی با استفاده از تکنیک لایه‌های کانولوشن دو بُعدی به داده‌های سه‌بُعدی تبدیل‌می‌شوند این مزیت این امکان را به مدل پیشنهادی می‌دهد که از روش‌های تقسیم‌کننده شاخه‌ای و تکنیک جایگشت استفاده‌کند تا عملکرد را بهبود بخشد. همچنین، تکنیک جایگشت در مدل پیشنهادی باعث ترکیب موازی لایه‌های کانولوشنی و لایه‌های حافظه طولانی کوتاه‌مدت می‌شود که از مزایای هر دو نوع لایه استفاده‌می‌کند. لایه‌های کانولوشنی بیشتر برای استخراج ویژگی‌های مکانی و لایه‌های حافظه طولانی کوتاه‌مدت بیشتر برای استخراج ویژگی‌های زمانی استفاده‌می‌شوند. در این روش، ابتدا داده‌های یک بُعدی به داده‌های دوبُعدی **n×n** تبدیل می‌شوند و سپس با استفاده از لایه‌های شبکه کانولوشنی دوبُعدی با اندازه هسته یک، داده‌ها به داده‌های سه‌بُعدی تبدیل می‌شوند. در مرحله بعدی، از جداکننده کانال‌ها و تکنیک جایگشت برای ایجاد دو شاخه متفاوت استفاده‌می‌شود و سپس داده‌های سه‌بُعدی توسط دو شاخه متفاوت از لایه‌های شبکه کانولوشنی و لایه‌های شبکه عصبی بازگشتی استخراج ویژگی می‌شوند.

مدل پیشنهادی **CLBS3** دارای ویژگی‌های زیر است:

- تبدیل داده‌ها از یک بُعد به دو بُعد و سپس سه بُعد بدون از دست رفتن ویژگی‌های پایه
- پیچیدگی محاسباتی کم به نسبت داده‌های سه‌بُعدی عمومی
- دقت تشخیص بالا با توجه به ترکیب لایه‌های بازگشتی و کانولوشنی
- دقت تشخیص بالا با توجه به داده های سه‌بُعدی
- استفاده از تکنیک‌های جداکننده کانال‌ها و تکنیک جایگشت برای ایجاد دو شاخه متفاوت
- استفاده از لایه‌های کانولوشنی برای استخراج ویژگی‌های مکانی
- استفاده از لایه‌های بازگشتی برای استخراج ویژگی‌های زمانی

شبکه‌های کانولوشن سه‌بُعدی برای تشخیص الگوها در داده‌های سه‌بُعدی استفاده‌می‌شوند. این شبکه‌ها با استفاده از عملیات کانولوشن، تشخیص و استخراج ویژگی‌های مهم را از داده‌های سه‌بُعدی انجام می‌دهند. دقت این شبکه‌ها نیز برای تشخیص الگوها و ویژگی‌های مختلف در داده‌های سه بُعدی بسیار بالا است ازاین‌رو تغییر دیدگاه شبکه عصبی و تبدیل ورودی داده‌های شبکه به بردارهای سه‌بُعدی می‌تواند دقت طبقه‌بندی را افزایش دهد. به‌عنوان‌مثال، در یک مدل با ۶۴ ویژگی در ورودی، در صورت نمایش به‌صورت برداری، هر ویژگی حداکثر دارای دو همسایه است. در نمایش دوبعدی این تعداد همسایگان به ۸ می‌رسد. با استفاده از مثلاً فیلتر ۳ × ۳ × ۳، فرمت ورودی می‌تواند به سه‌بُعدی تغییر یابد و تعداد همسایگان به ۲۶ برسد. با افزایش تعداد همسایگان، اطلاعات بیشتری در دسترس شبکه قرارمی‌گیرد که همین امر باعث‌می‌شود دقت شبکه افزایش یابد. بدیهی است که زمان آموزش و پیچیدگی نیز افزایش خواهدیافت. اگرچه امکان استفاده مستقیم از فیلترهای همگرا در ایجاد داده‌های سه‌بُعدی وجوددارد، اما در روش CLBS3 از یک تکنیک ساده بر اساس دو بُعدی برای ایجاد داده‌های سه بُعدی استفاده‌شده است. این روش نسبت به روش سه بُعدی معمولی پیچیدگی کمتری دارد.

در مدلCLBS3 ، به‌جای حرکت‌دادن بردار ورودی شبکه در فضای سه‌بعدی و استفاده از فیلتر کانولوشن سه بُعدی بر روی آن، ابتدا یک‌لایه انتقال (یک لایه کانوولوشن دوبعدی با اندازه فیلتر یک) اعمال‌می‌شود. ورودی به بردار سه‌بُعدی تبدیل‌می‌شود. این مدل از دو روش جابجایی برای استخراج اطلاعات همبستگی مفید استفاده‌می‌کند. اگر ورودی مدل S×L باشد، با استفاده از روش پیشنهادی، ورودی به ابعاد $S\times\sqrt{([L])}$

([L])√×تغییرمی‌کند که S تعداد رکوردها و L تعداد ویژگی‌های ورودی است. شکل ۴ روش اولین جابجایی را نشان‌می‌دهد. به‌عنوان مثال، مجموعه داده UNSWNB15 حاوی ۴۳ ویژگی است. با استفاده از مدل پیشنهادی تبدیل به ورودی[۴۳ √] ×[۴۳ √] خواهدشد؛ بنابراین، ۴۳ ویژگی اصلی تبدیل به ۴۹ ویژگی می‌شود که ۶ ویژگی جدید اضافه‌شده با صفر مقداردهی می‌شوند. این مدل شبکه کانولوشنی، سطرهای مربوط به هر کانال ورودی را می‌گیرد، آن‌ها را حرکت‌می‌دهد و آن‌ها را به‌عنوان یک کانال واحد در خروجی قرار می‌دهد؛ بنابراین، ورودی شبکه یک ماتریس ۷×۷×۲ است و خروجی آن ماتریس ۷×۲×۷ است که در شکل ۴ نشان داده‌شده‌است. نوع دوم جابه‌جایی در شکل ۵ نشان داده‌شده‌است. این نوع جابه‌جایی یک ورودی ۷×۷×۲ را، با تغییر مکان ستون‌های متناظر با هر کانال ورودی و قراردادن آن‌ها به‌عنوان یک کانال به ۲×۷×۷ تبدیل‌می‌کند. این دو نوع نگاشت باعث نزدیک شدن ویژگی‌های دوردست در فضای کانال می‌شوند و باعث‌می‌شود اطلاعات همبستگی از پارامترهای دوردست در داده‌های تغییریافته استخراج‌شود.

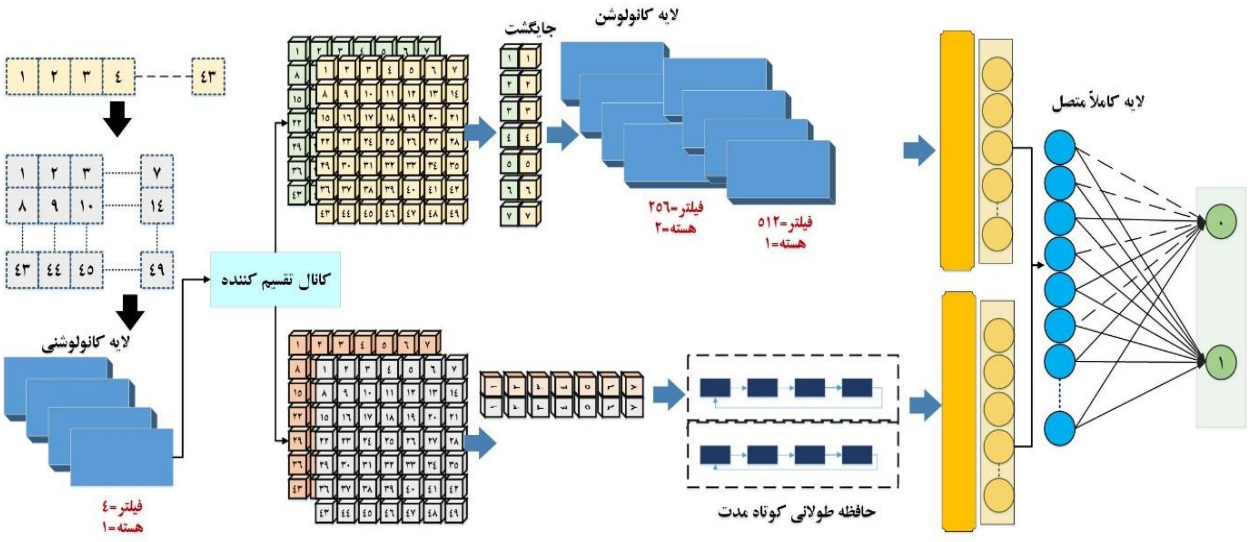
شکل ۴ نشان‌دهنده عملکرد مدل عمیق CLBS3 پیشنهادی بر روی مجموعه داده UNSWNB15 است که شامل قابلیت‌های زیر است:

۱-آماده‌سازی داده: یک نمایش دوبُعدی از بردار ورودی به‌عنوان ورودی مدل پیشنهادی استفاده‌می‌شود. لایه انتقال با چهار فیلتر کانولوشن به‌اندازه استاندارد ۱ اعمال‌می‌شود. ورودی به چهار کانال ارسال‌می‌شود که به دو بخش مساوی تقسیم‌شده و به دوشاخه مدل ما تغذیه می‌شود. کانال‌های آبی و نارنجی به شاخه دوم منجر می‌شوند درحالی‌که کانال‌های زرد و سبز به شاخه اول می‌روند. هر شاخه از یکی از انواع جابجایی مشخص‌شده برای استخراج داده‌های همبستگی مهم‌تر، استفاده‌می‌کند.

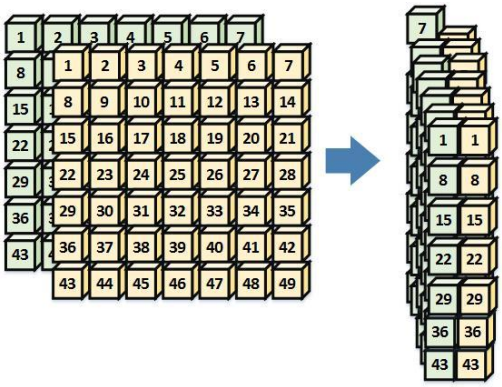
۲- شاخه اول: شاخه اول جابجایی خطی را بر روی ورودی انجام می‌دهد و به بردار ورودی ۷ × ۲ × ۷ تبدیل‌می‌شود که به‌عنوان ورودی یک شبکه کانولوشنی در نظر گرفته‌می‌شود. شبکه کانولوشن جدید شامل دو لایه با یک فیلتر و ۲۵۶ کانال به‌اندازه ۲ و یک فیلتر به‌اندازه ۱ است. این شبکه شامل یک‌لایه تخته‌کننده و یک‌لایه کاملاً متصل به ۵۱۲ با تابع فعالیت RELUE است. تمرکز این شاخه بیشتر بر استخراج ویژگی‌های فضایی است. شبکه‌های کانولوشن در تشخیص ویژگی‌های محلی بهتر عمل‌می‌کنند.

۳- شاخه دوم: شاخه دوم از جابجایی ستونی شبکه کانولوشن دوبعدی استفاده‌می‌کند و ورودی را به یک ماتریس سه‌بعدی به‌اندازه ۲ × ۷ × ۷ تبدیل‌می‌کند. این ماتریس سه‌بعدی به‌عنوان ورودی شبکه حافظه طولانی کوتاه‌مدت در نظر گرفته‌می‌شود. خروجی شبکه حافظه طولانی کوتاه‌مدت یک‌لایه کاملاً متصل به‌اندازه ۵۱۲ با تابع فعالیت RELUE است. تمرکز این شاخه بیشتر بر استخراج ویژگی‌های زمانی است.

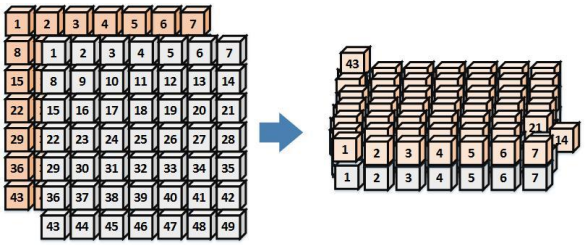
۴-لایه آخر: بر روی خروجی هر دو شاخه اعمال‌می‌شود و به‌عنوان یک‌لایه پیوندی عمل‌کرده و مسئول دسته‌بندی خروجی است؛ بنابراین، ورودی را به یک بردار یک‌بعدی که خروجی را نشان می‌دهد تبدیل‌می‌کند. تابع فعالیت SOFTMAX برای دسته‌بندی دودویی و تشخیص حمله استفاده‌می‌شود. در این لایه از بهینه‌ساز Adam و نرخ یادگیری ۰/۲ استفاده‌می‌شود.



شکل ۳: معماری مدل CLBS3



شکل ۴: جایگشت نوع اول



شکل ۵: جایگشت نوع دوم

۴. ارزیابی و نتایج

مدل CLBS3 پیشنهادی با پایتون پیاده‌سازی و ارزیابی‌شده است. آزمایش‌ها برای مقایسه عملکرد این مدل بر روی مجموعه داده UNSWNB15 و TONIOT انجام‌شده‌است. در این بخش ابتدا مجموعه داده‌های استفاده شده برای ارزیابی معرفی‌می شوند. سپس معیارهای ارزیابی موردا ستفاده معرفی‌می شود. در ادامه، از روش پیشنهادی برای تشخیص نفوذ در مجموعه داده تعریف‌شده با هدف افزایش دقت استفاده‌می‌شود. مقایسه‌ها با مدل‌های ترکیبی شامل مدل‌های شبکه‌های کانولوشنی[۲۱]، حافظه طولانی کوتاه‌مدت[۲۳]، کانولوشنی و حافظه طولانی کوتاه‌مدت[۲۷]، شبکه عصبی عمیق[۲۸]، و ترکیب شبکه‌های کانولوشنی و حافظه طولانی کوتاه‌مدت [۴۵] بر اساس تحقیقات قبلی صورت گرفته‌است .علاوه بر این، عملکرد شبکه کانولوشنی بهبودیافته جدید نیز به‌صورت جداگانه ارزیابی‌شده و با مدل یکپارچه پیشنهادی CLBS3 مقایسه‌می‌شود.

۴,۱ مجموعه داده‌ها

در این مقاله، برای ارزیابی عملکرد مدل پیشنهادی و الگوریتم‌های دیگر در مسئله تشخیص نفوذ شبکه‌های اینترنت اشیا، دو مجموعه داده معتبر مورداستفاده قرارمی‌گیرد. این مجموعه داده‌ها دارای ویژگی‌های منحصربه‌فرد برای تشخیص حملات سایبری در شبکه‌های اینترنت اشیا می‌باشند. این دو مجموعه داده شامل ToN-IoT، UNSW-NB 15 و هستند که تعداد نمونه‌های آن‌ها در جدول ۲ نشان داده‌شده‌است. در ادامه، توضیحات مختصری در مورد هر مجموعه داده ذکرشده‌است.

جدول ۲: مشخصات مجموعه داده‌ها

مجموعه داده	UNSW-NB15	ToN-IoT
تعداد رکوردهای نرمال	۱۶۴۶۷۳	۳۰۰۰۰۰
تعداد رکوردهای حمله	۹۳۰۰۰	۱۶۱۰۱۴۳
تعداد کل ویژگی‌ها	۴۵	۴۲
تعداد کلاس طبقه‌بندی	۲	۲

- **مجموعه داده ToN-IoT**

تمام نمونه‌های داده در مجموعه داده ToN-IoT از ۴۵ ویژگی تشکیل‌شده‌است. ۴۳ ویژگی اول برای صفت‌ها استفاده‌می‌شود و ویژگی ۴۴ ام با عنوان عادی بدون حمله و یا حمله، برچسب‌گذاری‌می‌شود و ویژگی آخر با عنوان بدون حمله یا نوع حمله‌ها برچسب‌گذاری شده‌است. مجموعه داده‌ای که برای آموزش و تست استفاده‌شده‌است شامل ۴۶۱۰۴۳ رکورد داده شبکه است. ۳۰۰۰۰۰ رکورد از آن‌ها عادی است و ۱۶۱۰۴۳ رکورد، بیانگر حمله است.

- **مجموعه داده UNSW-NB15**

مجموعه داده UNSW-NB15 که در آزمایشگاه امنیت سایبری استرالیا ایجادشده، برای تولید و شبیه‌سازی داده‌های واقعی و به‌روز طراحی‌شده‌است. این مجموعه داده شامل ۴۹ ویژگی است که توسط دوازده الگوریتم برای ایجاد برچسب‌های کلاس تولیدشده‌اند. بخشی از این داده‌ها به دو بخش مجزا تقسیم‌می‌شود: یک بخش برای آموزش با ۱۷۵۳۴۱ رکورد و دیگری برای تست با ۸۲۳۳۲ رکورد که هر دو شامل نمونه‌هایی از رفتارهای معمولی و انواع مختلف حملات هستند. هر مجموعه داده دارای ۴۵ ویژگی است. از این ۴۵ ویژگی، ۴۳ ویژگی برای توصیف صفات به کار می‌روند، ویژگی ۴۴ به‌عنوان نرمال یا حمله برچسب‌گذاری‌می‌شود و آخرین ویژگی نیز نوع نرمال یا حمله را مشخص‌می‌کند.

۴,۲ معیارهای ارزیابی

در این بخش، معیارهای مختلف برای ارزیابی عملکرد مدل پیشنهادی و سایر الگوریتم‌های مقایسه‌ای معرفی‌می‌شود. با توجه به مسئله طبقه‌بندی، چهار معیار از رایج‌ترین معیارهای اعتبارسنجی شامل دقت، امتیاز f1، نرخ مثبت کاذب[۲] و نرخ منفی کاذب[۳] با ماتریس درهم‌ریختگی قابل‌تعریف است.

بر اساس ماتریس درهم‌ریختگی، مثبت واقعی[۴] نشان‌دهنده تعداد نمونه‌های ترافیک حملاتی است که به‌درستی توسط مدل پیشنهادی طبقه‌بندی‌شده‌اند. منفی واقعی[۵] نمونه ترافیک عادی را نشان می‌دهد که به‌درستی توسط مدل پیشنهادی طبقه‌بندی‌شده‌اند. مثبت کاذب تعداد نمونه ترافیک‌های عادی هستند، اما به‌اشتباه توسط مدل پیشنهادی به‌عنوان نمونه حمله طبقه‌بندی‌شده‌اند. منفی کاذب[۶] تعداد نمونه ترافیک‌ها حملاتی هستند، اما به‌اشتباه توسط مدل پیشنهادی به‌عنوان نمونه ترافیک عادی طبقه‌بندی‌شده‌اند.

$$\text{Accurecy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{F1} - \text{score} = 2 * \frac{precision * recall}{precision + recall}$$

$$FPR = \frac{FP}{FP + TN}$$

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

۴،۳ تنظیم پارامترها

در مدل یادگیری عمیق، هایپرپارامترها مواردی هستند که تاثیر بسیار زیادی بر عملکرد و دقت مدل دارند. تنظیم و بهینه‌سازی هایپرپارامترها بسیار مهم است و برای به‌دست آوردن یک مدل یادگیری عمیق با دقت بالا، نیاز است که این هایپرپارامترها به‌درستی تنظیم‌شوند. با تغییر پارامترهای مهم، عملکرد سیستم تشخیص نفوذ به‌طور قابل‌توجهی تغییرمی‌کند. تأثیر تغییر هر یک از پارامترهای مهم به‌طور خلاصه در این بخش نشان داده شده است. می‌توان هایپر پارامترهای دخیل را در دو بخش مجزا توضیح‌داد.

- بهینه‌سازی هایپرپارامترهای مدل پیشنهادی با استفاده از Optuna

در این مقاله، برای بهبود عملکرد مدل شبکه عصبی کانولوشنی، از بهینه‌سازی هایپرپارامترها استفاده‌شده‌است. مدل ما روی داده‌های UNSW-NB15 آموزش داده‌شده و ابزار Optuna برای یافتن بهترین مقادیر هایپرپارامترها به‌کاررفته‌است Optuna .از روش بهینه‌سازی بیزی استفاده‌می‌کند که به‌طور کارآمد مقادیر مختلف هایپرپارامترها را بررسی‌کرده و بهترین ترکیب را برای رسیدن به حداکثر دقت انتخاب‌می‌کند.

Optuna از یک فرآیند هوشمندانه بهینه‌سازی استفاده‌می‌کند که در آن به‌جای بررسی تمام ترکیبات ممکن، ترکیب‌های امیدوارکننده را براساس نتایج قبلی انتخاب‌می‌کند. در این روش، Optunaبا استفاده از توزیع احتمالات بیزی و یادگیری از نتایج هر آزمایش، ترکیب‌های جدیدی از هایپرپارامترها را پیش‌بینی می‌کند که احتمالاً دقت بالاتری دارند. نحوه عملکرد optuna به‌صورت زیر است:

۱-آغاز با نمونه‌های تصادفی :در ابتدای کار، Optunaبا بررسی تعدادی ترکیب تصادفی از هایپرپارامترها، نتایج اولیه را جمع‌آوری‌می‌کند.

۲-مدل‌سازی توزیع احتمالاتی بیزی :پس از دریافت نتایج اولیه، Optunaبر اساس مدل بیزی، توزیعی از احتمال موفقیت هر ترکیب را بر اساس نتایج قبلی ایجادمی‌کند.

۳-انتخاب ترکیبات جدید با جستجوی کارآمد Optuna :با استفاده از الگوریتم‌های جستجو، ترکیب‌های جدیدی را از این توزیع احتمالاتی انتخاب‌می‌کند که شانس بیشتری برای بهبود دقت مدل دارند.

۴-تکرار و بهبود تدریجی :در هر مرحله از جستجو، نتایج به مدل احتمالاتی اضافه‌می‌شود و Optuna به‌تدریج به ترکیب‌هایی نزدیک می‌شود که بالاترین دقت را دارند.

این رویکرد بیزی باعث‌می‌شود که Optuna بدون نیاز به آزمون همه ترکیبات، بهینه‌ترین مقادیر را پیداکند و از نظر زمان و منابع محاسباتی بسیار کارآمد باشد.

در مدل پیشنهادی، شش هایپرپارامتر اصلی زیر برای بهینه‌سازی در نظر گرفته‌شدند. بازه‌های تعریف‌شده برای هر یک از هایپرپارامترها بر اساس تجربیات مقالات قبلی و تجربه است. این هایپرپارامترها به‌طور مستقیم بر دقت و کیفیت یادگیری مدل تأثیر می‌گذارند:

۱-تعداد فیلترها در لایه کانولوشن :این پارامتر مشخص‌می‌کند که چند ویژگی از ورودی توسط لایه کانولوشن استخراج‌شود. برای این پارامتر، بازه‌ای بین ۳۲ تا ۶۴ انتخاب شده‌است. افزایش تعداد فیلترها می‌تواند مدل را قادر به یادگیری ویژگی‌های پیچیده‌تر کند.

۲-اندازه‌ی فیلتر این پارامتر اندازه هر فیلتر را مشخص‌می‌کند. برای اندازه فیلتر بازه‌ای بین ۳ تا ۵ در نظر گرفته‌شد که اندازه‌های کوچک‌تر باعث استخراج ویژگی‌های جزئی‌تر و اندازه‌های بزرگ‌تر باعث تعمیم‌دهی بیشتر می‌شوند.

۳-تعداد گره‌ها در لایه :Dense این پارامتر نشان‌دهنده تعداد گره‌های لایه چگال (Dense) است که می‌تواند قدرت یادگیری مدل را در مراحل پایانی شبکه افزایش دهد. بازه انتخابی بین ۶۴ تا ۱۰۲۴ تنظیم‌شده‌است.

۴-نرخ :Dropout روشی برای جلوگیری از بیش برازش است که باعث می‌شود برخی از گره‌ها به‌طور تصادفی در زمان آموزش غیرفعال شوند. این نرخ در بازه ۰،۱ تا ۰،۵ انتخاب شد که مقادیر بالاتر باعث کاهش بیش برازشی می‌شود.

۵-نرخ یادگیری: نرخ یادگیری میزان تغییرات وزن‌ها را در هر به‌روزرسانی مدل تعیین‌می‌کند. برای این پارامتر از بازه‌ی ۰،۰۰۰۱ تا ۰،۱ به‌صورت لگاریتمی استفاده‌شد که مقدار بالاتر منجر به یادگیری سریع‌تر و مقادیر کمتر منجر به یادگیری آهسته‌تر ولی پایدارتر می‌شوند.

۶-انتخاب بهینه‌ساز: برای بهینه‌سازی وزن‌ها، دو الگوریتم پرکاربرد SGDو Adamدر نظر گرفته‌شده‌اند.

در این بهینه‌سازی، تعداد تکرارها برای Optuna برابر با ۲۰ تنظیم‌شده است؛ یعنی ابزار Optuna صد ترکیب مختلف از مقادیر هایپرپارامترها را به‌طور هوشمندانه انتخاب و آزمایش می‌کند تا بهترین نتایج به دست آید. همچنین، برای هر ترکیب از هایپرپارامترها، مدل به مدت ۵۰ دوره (epoch) روی داده‌های آموزشی آموزش داده‌می‌شود. این تعداد ازepoch ها اجازه‌می‌دهد که مدل به‌خوبی آموزش ببیند، اما از طرفی باعث افزایش زمان اجرای برنامه نیز می‌شود.

Optunaپس از آزمایش ۲۰ترکیب مختلف از مقادیر هایپرپارامترها، ترکیبی از مقادیر را به عنوان بهترین مجموعه بهینه‌یافته پیشنهادمی‌دهد. این ترکیب با بیشترین دقت ممکن در داده‌های آزمایشی به‌عنوان بهترین حالت انتخاب‌می‌شود

. جدول ۳، تأثیر هر یک از هایپرپارامترها را روی نتیجه نهایی مدل نشان‌می‌دهد با استفاده از بهینه‌سازی بیزی درOptuna ، به مجموعه‌ای از مقادیر هایپرپارامترها دست‌یافتیم که بالاترین دقت ممکن را برای مدل ما فراهم‌کردند. نتایج نشان‌دهنده موفقیت این روش در تعیین مقادیر بهینه و همچنین ارزشمندی آن در جلوگیری از آزمایش‌های دستی و زمان‌بر هایپرپارامترها است. بهترین دقت برای زمانی است که تعداد ۴۵ فیلتر به اندازه ۳ در شبکه کانولوشنی داشته‌باشیم و تعداد گره‌های لایه dense برابر ۱۲۷ باشد. با بهینه‌ساز adam و نرخ یادگیری ۰.۰۰۰۸۲/ ا بیشترین دقت برابر۰۹۷/۳درصد و کمترین زمان برابر ۱۵۸

ثانیه به‌دست آمده‌است. بنابراین برای اجرای مدل هایپرپارامترها با این مقادیر تنظیم‌شده‌اند. .

جدول ۳: تنظیم هایپرپارامترها و زمان

id	فیلتر	اندازه فیلتر	گره	dropout	نرخ یادگیری	بهینه ساز	دقت	زمان
1	52	5	840	0.410	0.0485	adam	0.68	212
2	50	5	648	0.286	0.0193	sgd	0.958	192
3	49	5	193	0.1472	0.0091	adm	0.969	183
4	57	5	508	0.160	0.00024	sgd	0.844	207
5	35	5	397	0.3887	0.00034	sgd	0.846	208
6	41	3	341	0.286	0.0016	sgd	0.922	201
7	42	5	462	0.2077	0.00155	sgd	0.927	203
8	40	3	393	0.12581	0.00025	adm	0.968	188
9	36	3	75	0.3911	0.04161	adm	0.951	190
10	63	4	293	0.3083	0.0609	sgd	0.961	187
11	45	3	127	0.187	0.00082	adm	0.973	158
12	60	5	208	0.1336	0.0001	sgd	0.759	210
13	55	3	173	0.238	0.0028	adm	0.971	162
14	52	4	71	0.454	0.00811	adm	0.967	184
15	50	4	188	0.228	0.0094	adm	0.969	183
16	55	4	186	0.2503	0.0076	adm	0.969	182
17	47	3	145	0.2080	0.023	adm	0.962	180
18	48	4	238	0.3149	0.0027	adm	0.971	160
19	45	3	257	0.333	0.00169	adm	0.970	162
20	53	4	262	0.35957	0.007	adm	0.968	186

- **هایپر پارامترهای اجرا**

برای اعتبارسنجی مجموعه داده‌ها، اعتبارسنجی متقابل k-fold انجام شده است. بهترین عملکرد در پنجاه درصد داده‌های آموزشی و پنجاه درصد داده‌های آزمایشی به‌دست آمده‌است. حداقل تعداد تکرار تنظیم-شده در هر اجرا برابر با ۱۰۰ در نظر گرفته شده‌است.

در حین اجرای روشهای یادگیری، وجود پارامترهای تصادفی باعث تولید مقادیر متفاوت و متغیر شده‌است؛ بنابراین، برای هر مورد، اجراها به صورت ۱۰ بار (هر بار به مدت ۱۰۰ دور) انجام شده‌است و میانگین نتایج این اجراها به عنوان جواب نهایی در نظر گرفته‌می‌شود.

۴,۴ نتایج ارزیابی

اطلاعات اولیه ممکن ا ست حاوی اطلاعات نامعلوم و همراه نویز با شند که بدون شک بر کارایی مدل تأثیرگذار خواهدداشت. مدل ارائه‌شده نیازمند به‌کارگیری چندین مرحله پیش‌پردازش قبل از استفاده از تکنیک‌های CLBS3 است. این فرآیندهای پیش‌پردازش به بهبود کیفیت داده‌های مربوط به اینترنت ا شیا، افزایش دقت مدل ارائه شده و کاهش پیچیدگی محاسبات کمک‌می‌کنند. سه مرحله اصلی پیش‌پردازش شامل پاک‌سازی، رمزنگاری و نرمال سازی داده‌ها هستند که در بخش ۳,۱ شرح داده‌شده‌اند. در مرحله نرمال‌سازی از تکنیک‌های مختلف نرمال سازی ا ستفاده شده‌ا ست. جدول ۴ مقادیر نرمال سازی داده ها برای تکنیک های نر مال‌سازی مختلف را نشان‌می‌د هد. تکنیک‌های نرمال‌سازی مورد استفاده شامل نرمال‌سازی حداقل-حداکثر، نرمال‌سازی z-score و نرمال‌سازی تانژانت هیپربولیک هستند. نتایج نشان‌می‌دهد که روش نرمال‌سازی حداقل-حداکثر دقت بهتری را به د ست آورده‌ا ست؛ بنابراین، در مرحله پیش‌پردازش از روش حداقل-حداکثر برای نرمال‌سازی داده‌ها استفاده‌شده‌است.

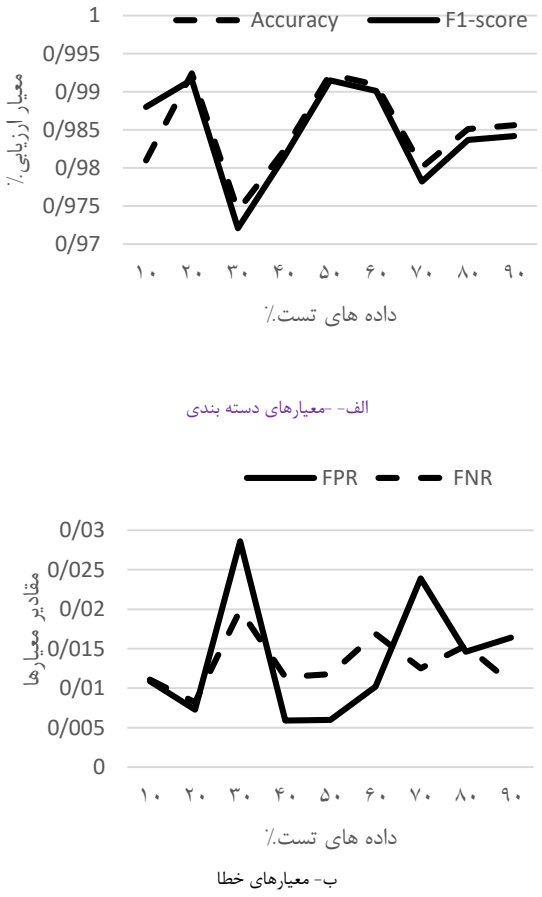
جدول ۴: نرمال‌سازی داده‌ها با تکنیک‌های نرمال‌سازی مختلف

مجموعه داده	تکنیک های نرمال‌سازی	مجموعه داده		
		دقت	صحت	امتیاز f1
TONIOT	Min-max	۹۴/۸	۹۲/۱۵	۹۳/۲
	z-score	۸۵/۸	۸۳/۹۶	۸۴/۱۲
	tanh	۹۳/۲	۸۸/۱۳	۹۲/۹
UNSWNB15	Min-max	۹۲/۹۷	۹۳/۱۵	۹۳/۱
	z-score	۹۱/۲۵	۹۲/۴	۹۳
	tanh	۸۹/۳	۹۰	۹۳/۴

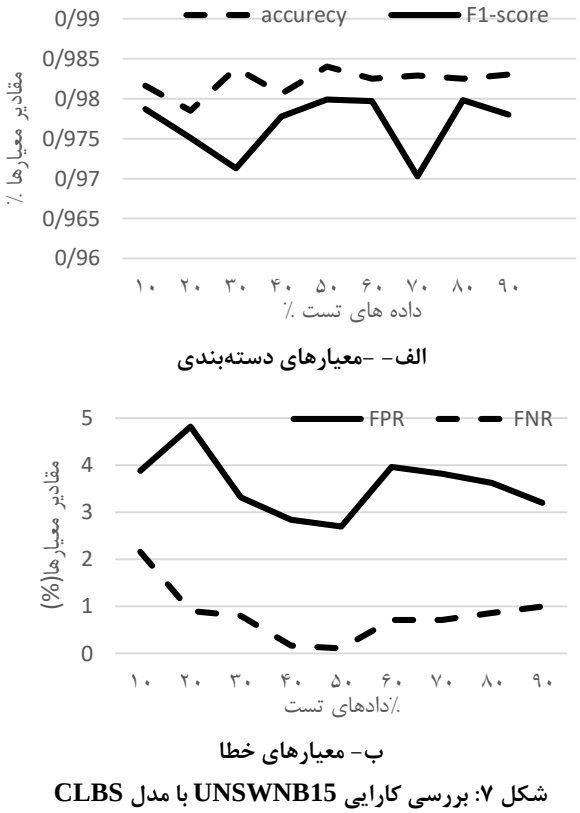
در ادامه این بخش مدل پیشنهادی CLBS3 بر روی مجموعه داده UNSWNB15 و TONIOT ارزیابی‌می‌شود. ذو معیار اصلی دسته‌بندی شامل دقت و امتیاز f1 و همچنین دو معیار مهم دیگر یعنی FPR و FNR برای مقایسه مدل پیشنهادی با سایر روش‌ها در نظر گرفته‌شده‌است. از تکنیک اعتبارسنجی متقابل k-fold برای اعتبارسنجی مدل پیشنهادی و سایر مدل‌ها استفاده‌شده‌است تا نشان داده‌شودکه مدل پیشنهادی تا چه اندازه در داده‌های تست متفاوت، نتایج پایداری دارد. نتایج آزمایش تکنیک اعتبارسنجی K-Fold بر روی مجموعه داده UNSWNB15 نشان‌می‌دهد که مدل پیشنهادی در ۱۰ تکرار مختلف، در ۵۰ درصد داده‌های آزمایشی عملکرد بهتری دارد. بدین ترتیب نتایج اعتبارسنجی روی مجموعه داده‌های IToN-IoT، UNSWNB15، به ترتیب در شکل ۶ و ۷ آورده شده‌است

شکل ۸ ماتریس درهم ریختگی روش CLBS3 را در مجموعه داده UNSWNB15 نشان‌می‌دهد. ماتریس درهم‌ریختگی گویای دقت بالا در تشخیص است. با استفاده از این ماتریس، می‌توان به‌راحتی دقت الگوریتم در تشخیص حمله‌ها را محاسبه‌کرد و عملکرد الگوریتم را ارزیابی‌کرد.

در این ماتریس، مثبت واقعی نشان‌دهنده تعداد مواردی است که الگوریتم به‌درستی حمله را تشخیص داده‌است که ۵۹۲۱۴ نمونه است، منفی کاذب نشان‌دهنده تعداد مواردی است که حمله وجود داشته اما الگوریتم آن را به عنوان حمله تشخیص‌نداده‌است که ۴۶۹ نمونه است، مثبت کاذب نشان‌دهنده تعداد مواردی است که الگوریتم به اشتباه حالتی که حمله نبوده‌است را حمله تشخیص‌داده‌است که ۱۰۵۱ نمونه است و منفی درست نشان‌دهنده تعداد مواردی است که الگوریتم به-درستی عدم حمله را تشخیص‌داده‌است که ۲۶۹۳۷ مورد است. در روش-های مورد مقایسه با وجود اینکه دقت مذل بالا است ولی تعداد رکوردهای منفی کاذب خیلی بیشتر است در خالی که در حالت حمله



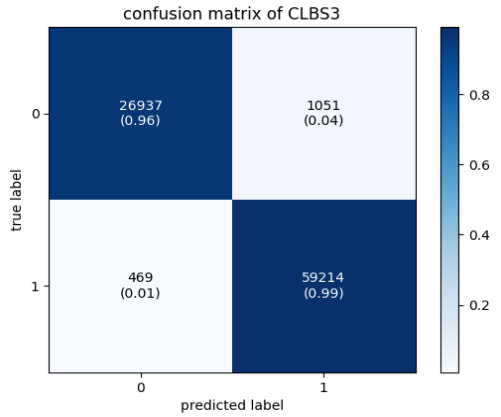
شکل ۶: بررسی کارایی TONIOT با مدل CLBS



شکل ۷: بررسی کارایی UNSWNB15 با مدل CLBS

هر چقدر حملات به درستی تشخیص‌داده‌شوند نشان‌دهنده قوی بودن مدل است. با تحلیل ماتریس درهمریختگی می‌توان نتیجه‌گرفت که مقدار کم FPR در الگوریتم پیشنهادی باعث دقت بالا در دسته‌بندی دودویی می‌شود؛ بنابراین، می‌توان‌گفت که مدل پیشنهادی بهتر از مدل‌های دیگر برای مجموعه داده‌های مورد بررسی عمل‌می‌کند

.



شکل ۸: ماتریس درهمریختگی مجموعه داده UNSWNB15

برای مقایسه روش پیشنهادی، جدول ۵ نتایج ارزیابی روش پیشنهادی را بر روی مجموعه داده UNSWNB-15 نسبت به مدل‌های شبکه‌های کانولوشنی[۲۱]، حافظه طولانی کوتاه‌مدت[۲۳]، کانولوشنی و حافظه طولانی کوتاه‌مدت[۲۷]، شبکه عصبی عمیق[۲۸]، و ترکیب شبکه‌های کانولوشنی و حافظه طولانی کوتاه‌مدت [۴۵] نشان‌می‌دهد.

در انتخاب الگوریتم‌های مورد مقایسه سعی‌شده‌است، الگوریتم هایی انتخاب‌شوند که از ترکیب شبکه‌های کانولوشنی و خافظه کوتاه بلندمدت استفاده‌کنند و خود روش‌های مورد مقایسه نیز در تحقیقاتشان با روش-های قدیمی بیشتری مقایسه‌شده‌باشند. از جدول ۵ مشخص است که روش پیشنهادی عملکرد قابل مقایسه‌ای نسبت به مدل‌های موجود دارد. تمام مدل‌های مقایسه‌شده دقت تقریباً بالایی دارند ولی معمولاً توازن بین دقت و امتیاز F1 وجودندارد. به عنوان نمونه مطالعه [۲۸]دارای دقت بی‌شتر ولی FPR بیشتر نیز هست یا مطالعه [۲۳] توازنی بین دقت و امتیاز F1 وجودندارد. مدل پیشنهادی بالاترین دقت و اندازه‌گیری F را دارد و در عین حال مقادیر کمی از FNR را نیز دارد.

جدول ۵: مقایسه مدل پسشنهادی با روش‌های دیگر در مجموعه داده UNSWNB15

مرجع	مدل	دقت	امتیاز F1	FPR	FNR
[۲۱]	CNN	۹۲/۸۵	۹۴/۲۵	۱۷/۸	۱۷/۴۵
[۲۳]	LSTM	۹۳	۹۶/۳۷	۱۱/۵۵	۷
[۲۷]	CONV-LSTM	۹۴/۷۵	۹۴/۷۸	۹/۸	۴/۴۵
[۲۸]	DNN	۹۵	۹۵/۶۷	۱۱/۷	۶/۶۷
[۴۵]	OCNN	۹۴/۶۷	۹۷/۵۱	۵/۰۳	۴/۱۲
	HMLSTM	۹۶/۲۲	۹۸/۰۵	۶/۷۵	۳/۶۹
	OCNN-HMLSTM CLBS3	۹۶/۳۳	۹۸/۱۳۲	۵/۸۷	۳/۶۷
روش پیشنهادی		۹۸/۴	۹۷/۹۹	۳/۷	۰/۷۹

مقایسه مدل پیشنهادی CLBS3 با مدل‌های پیشرفته دیگر بر روی مجموعه داده TON-IOT در جدول ۶ نشان‌می‌دهد که مدل پیشنهادی در معیار دقت نسبت به مدل‌های دیگر عملکرد بهتری دست‌یافته‌است. البته در معیارهای دیگر عملکرد الگوریتم پیشنهادی به‌طور میانگین حدود ۹۸ است که عملکرد بالایی الگوریتم پیشنهادی را بروی مجموعه داده TON-IOT ثابت‌می‌کند. چندین روش اخیر نیز بر روی مجموعه داده UNSWNB15 در جدول ۶ آمده‌است. نتایج ارزیابی، برتری روش پیشنهادی را در چهار معیار ارزیابی مهم نشان‌می‌دهد.

جدول (۶): مقایسه مدل پیشنهادی DLBS-3D با مدل‌های پیشرفته دیگر بر روی مجموعه داده TON-IOT

مرجع	الگوریتم	دقت	صحت	یادآوری	امتیاز F1
[۴۶]	DT	۸۸/۶۷	۸۴/۸۲	۸۲/۲۹	۸۳/۵۴
	NB	۷۴/۱۳	۶۰/۹۸	۷۱/۴۴	۶۵/۷۷
	RF	۹۳/۹۴	۸۶/۸۸	۹۷/۳۲	۹۱/۸۱
	Ensemble	۹۶/۳۵	۹۰/۵۴	۹۹/۹۸	۹۵/۰۳
[۴۷]	DNRANN-1	۹۹/۱۵	۹۹/۲۳	۹۹/۰۷	۹۹/۲۷
	DNRANN-2	۹۹/۲۰	۹۸/۹۱	۹۸/۸۶	۹۸/۸۸
	DNRANN-3	۹۸/۱۷	۹۸/۸۹	۹۸/۸۳	۹۸/۸۶
[۴۸]	DFF+AE	۹۶/۹۳	-	۹۸/۲۵	۹۸
	CNN+LDA	۹۷/۶۰	-	۹۹/۴۶	۹۹
	RNN+PCA	۹۶/۱۳	-	۹۶/۹۰	۹۸
	LR+LDA	۹۷/۶۸	-	۹۹/۵۹	۹۹
	DT+AE	۹۸/۲۳	-	۹۸/۲۸	۹۹
	NB+PCA	۹۷/۹۴	-	۹۹/۸۲	۹۹
[۴۹]	LSTM-400	۹۶/۵۲	-	۹۸/۳	۹۷/۳
	LSTM-500	۹۶/۵۶	-	۹۷/۳	۹۷/۳۵
CLBS3 (روش پیشنهادی)		۹۸/۲	۹۸/۲۱	۹۸/۲۶	۹۸/۲۷

جدول (۷): مقایسه مدل پیشنهادی CLBS3 با مدل‌های پیشرفته دیگر بر روی مجموعه داده UNSWNB15

مرجع	الگوریتم	دقت	صحت	یادآوری	امتیاز F1
[۵۰]	GWO-PSO-RF	۸۴/۴۶	۸۰/۸۵	۷۹/۴۳	۷۹/۰۱
	MOPSO-Levy-KNN	۸۷/۱۲	۸۵/۷۸	۸۶/۶۵	۸۷/۳۴
	AAFSA with GA-FR	۹۴/۴۸	۹۴/۲۹	۹۴/۲۶	۹۴/۴۲
[۴۸]	DFF+LDA	۹۸/۳۴	-	۹۹/۸۸	۸۵/۰
	CNN+LDA	۹۸/۲۸	-	۹۹/۸۹	۸۵
	RNN+ LDA	۹۸/۳۱	-	۹۸/۸۸	۸۵
	LR+FULL	۹۸/۴۷	-	۹۸/۸۸	۸۶
	DT+ FULL	۹۹/۲۷	-	۹۱/۵۸	۹۲
	NB+ LDA	۹۸/۳۴	-	۹۸/۳۹	۸۵
[۵۱]	LSTM	۹۱/۱۴	-	-	-
	CNN	۹۰/۰۹	-	-	-
	CNN LSTM	۹۳/۲۱	-	-	-
[۵۲]	Simple RNN	۸۷/۰۷	-	-	-
	LSTM	۸۵/۰۸	-	-	-
	GRU	۸۸/۴۲	-	-	-
روش پیشنهادی	CLBS3	**۹۸/۴**	**۹۷/۸**	۹۸/۷۸	**۹۷/۹۹**

۵. تحلیل مدیریت و نرخ خطا

در این مقاله نتایج به‌دست‌آمده از ارزیابی مدل پیشنهادی حاکی از کارایی و دقت آن در شناسایی حملات و تفکیک آن‌ها از ترافیک عادی است. مقادیر به‌دست‌آمده برای نرخ مثبت کاذب نرخ منفی کاذب به ترتیب برابر با ٪۳,۷۶ و ٪۰,۷۹ هستند.

نرخ مثبت کاذب به درصد مواردی اشاره‌دارد که مدل به‌اشتباه ترافیک عادی را به‌عنوان حمله شناسایی‌کرده‌است. مقدار ٪۳,۷۶ برای FPRنشان‌دهنده این است که از هر ۱۰۰ مورد ترافیک عادی، تقریباً ۳,۷۶ مورد به‌اشتباه به‌عنوان حمله شناسایی‌شده‌اند. این نرخ نسبتاً پایین به این معناست که مدل ما توانسته‌است ترافیک عادی را به‌طور مؤثری از حملات متمایزکند. کاهش تعداد هشدارهای کاذب نه‌تنها به کاهش بار بر روی گروه‌های امنیتی کمک‌می‌کند، بلکه موجب افزایش اعتبار سیستم نیز می‌شود. در دستگاه‌های امنیتی، وجود هشدارهای کاذب می‌تواند منجر به کاهش توجه و اعتبار گروه‌های امنیتی نسبت به هشدارها شود، بنابراین مدیریت این نرخ اهمیت زیادی دارد.

نرخ منفی کاذب به درصد حملات واقعی اشاره‌دارد که مدل قادر به شناسایی آن‌ها نبوده‌است. مقدار ٪۰,۷۹ برای نرخ منفی کاذب به این

معناست که از هر ۱۰۰ حمله واقعی، تنها حدود ۰,۷۹ مورد به‌اشتباه شناسایی‌نشده‌اند. این نرخ بسیار پایین نشان‌دهنده توانایی بالای مدل در شناسایی حملات واقعی است. در زمینه اینترنت اشیا، جایی که تهدیدات ممکن است جدی باشند، توانایی شناسایی صحیح حملات بسیار حیاتی است. یک نرخ منفی کاذب پایین به این معناست که سیستم می‌تواند به‌طور مؤثر و سریع به تهدیدات پاسخ‌دهد و از وقوع خسارات بیشتر جلوگیری‌کند.

مدیریت خطا در مدل‌های تشخیص نفوذ به فرایندهایی اطلاق‌می‌شود که هدف آن کاهش نرخ‌های مثبت کاذب و منفی کاذب است. در این مقاله، برای مدیریت خطا از تنظیم هایپرپارامترها و drouput استفاده‌شده‌است. از روش بیزین برای بهینه‌سازی هایپرپارامترهای مدل استفاده‌کردیم. این روش به ما این امکان را می‌دهد که ترکیب بهینه پارامترها را شناسایی‌کنیم. به عنوان مثال، با تنظیم مناسب نرخ یادگیری و تعداد لایه‌های مدل، توانسته‌ایم دقت آن را بهبودبخشیم و در نتیجه، نرخ‌های خطا را کاهش دهیم

برای جلوگیری از بیش برازش، از تکنیک Dropout بهره‌برداری کردیم. این تکنیک با حذف تصادفی برخی از نورون‌ها در هر دوره آموزش، به مدل کمک‌می‌کند تا به ویژگی‌های خاص وابسته نشود. این فرایند به بهبود قابلیت تعمیم‌پذیری مدل و کاهش نرخ منفی کاذب کمک‌کرده‌است.

نتایج نشان‌دهنده عملکرد مطلوب مدل در شناسایی حملات در محیط اینترنت اشیا است. با کاهش نرخ‌های مثبت و منفی کاذب، مدل ما توانسته‌است به بهینه‌سازی قابل‌توجهی دست یابد. با توجه به نتایج به‌دست‌آمده، می‌توان‌گفت که مدل ما می‌تواند به‌عنوان ابزاری مؤثر در تشخیص نفوذ در IoT مورداستفاده قرار گیرد و به تقویت امنیت دستگاه‌های IoT کمک‌کند.

۶. تحلیل مصرف منابع پردازشی و حافظه مدل پیشنهادی

این بخش برای ارزیابی کارایی مدل و اطمینان از مناسب بودن آن برای دستگاه‌های اینترنت اشیا (IoT) که معمولاً از منابع سخت‌افزاری محدودی برخوردارند، اهمیت زیادی دارد. نتایج آزمایش مصرف منابع پردازشی و حافظه مدل پیشنهادی بر روی مجموعه داده unswnb15 را می‌توان به‌صورت زیر دسته‌بندی‌کرد:

زمان اجرا: مدت زمان اجرای مدل برای پردازش داده‌ها برابر با ۸۱/. ثانیه اندازه‌گیری شد. این زمان اجرا نشان‌دهنده سرعت مناسبی برای مدل است که در کاربردهای IoT می‌تواند به معنای پردازش تقریبی در زمان واقعی یا نزدیک به آن باشد. زمان اجرای کم، از نظر بهره‌وری انرژی و پاسخ‌دهی سریع در دستگاه‌های IoT که باید به وقایع شبکه به‌سرعت واکنش نشان‌دهند، بسیار مطلوب است.

مصرف پردازنده :مصرف CPU در طی اجرای مدل برابر با تقریباً صفر درصد گزارش‌شد. این مقدار نشان‌می‌دهد که مدل، در شرایط فعلی اجرا، تقریباً هیچ بار پردازشی قابل‌توجهی بر CPU سیستم واردنمی‌کند. این می‌تواند به دلیل بهینه‌سازی‌های مدل و پردازش توسط بخش‌هایی از GPUیا استفاده از واحدهای محاسباتی دیگر باشد که بار پردازشی کمتری بر CPU دارند. برای دستگاه‌های IoT که اغلب پردازنده‌های با توان کم دارند، مصرف پایین CPU یک مزیت بزرگ محسوب‌می‌شود.

مصرف حافظه: مصرف حافظه در طول اجرای مدل به‌طور متوسط برابر با ۱۳/۷۰۳۴۰۷ مگابایت بود. این مقدار نسبتاً پایین نشان‌می‌دهد که مدل فضای حافظه زیادی را برای نگهداری پارامترها و محاسبات میانی اشغال‌نمی‌کند. چنین مصرف حافظه‌ای برای دستگاه‌های IoT که محدودیت‌های حافظه دارند، مزیت بزرگی است و نشان‌می‌دهد که مدل می‌تواند در چنین دستگاه‌هایی بدون نگرانی از مصرف بیش از حد حافظه، به‌راحتی اجراشود.

اوج مصرف حافظه: اوج مصرف حافظه مدل در طی اجرای آزمایش‌ها به ۴۷/۰۷۹۸۷ مگابایت رسید. این مقدار حداکثر حافظه‌ای است که مدل در زمان بارگیری و پردازش اطلاعات نیاز دارد. این افزایش لحظه‌ای در مصرف حافظه به دلیل ذخیره‌سازی موقت نتایج میانی در برخی از لایه‌های مدل، مانند لایه‌های کانولوشن و LSTM است. با این حال، اوج مصرف حافظه نیز همچنان در حد معقولی باقی‌مانده‌است و نشان‌می‌دهد که مدل در استفاده از حافظه بهینه عمل‌می‌کند.

مصرف RAM سیستم :میزان مصرف RAM کلی سیستم در طول اجرای مدل به ۲۰۵۸/۲۹۳۲۴۸ مگابایت رسید. این مقدار شامل حافظه‌ای است که سیستم‌عامل و سایر فرآیندها نیز اشغال‌کرده‌اند. این میزان مصرف نشان‌می‌دهد که مدل در زمان اجرا حجم معقولی از RAM را مصرف می‌کند و با توجه به محدودیت‌های سخت‌افزاریIoT ، می‌توان آن را برای دستگاه‌های IoT با ظرفیت حافظه مشابه مناسب در نظرگرفت.

این نتایج بیانگر آن است که مدل پیشنهادی برای کاربردهای IoT از نظر مصرف منابع پردازشی و حافظه بهینه طراحی‌شده و می‌تواند در محیط‌هایی با محدودیت‌های سخت‌افزاری اجراشود. این مدل ضمن بهره‌وری بالا، بار پردازشی کمی بر CPU دارد، حافظه زیادی اشغال نمی‌کند و زمان اجرای کوتاهی دارد که همگی ویژگی‌های مناسبی برای دستگاه‌های IoT به شمار می‌آیند.

۷. نتیجه‌گیری

در این مقاله یک رویکرد یادگیری عمیق سه‌بعدی جدید مبتنی بر تقسیم شاخه‌ای برای سیستم تشخیص نفوذ در شبکه‌های اینترنت اشیا به نام CLBS3 پیشنهادشد. در ابتدا داده‌های یک‌بعدی شبکه‌های اینترنت اشیا با تکنیک فیلتر لایه‌های کانولوشنی به داده‌های سه‌بعدی تبدیل‌شد و سپس از جداکننده کانال‌ها و تکنیک جایگشت برای ایجاد دو شاخه متفاوت و موازی استفاده شد. در مدل پیشنهادی CLBS3 از ترکیب موازی لایه‌های کانولوشنی و لایه‌های حافظه کوتاه‌مدت طولانی برای بهره‌برداری از مزایای هر دو مدل استفاده‌شده‌است .لایه‌های کانولوشنی برای استخراج ویژگی‌های مکانی به‌صورت سراسری و محلی

و لایه‌های حافظه کوتاه‌مدت طولانی بی شتر برای ا ستخراج ویژگی‌های زمانی یا به‌صـورت توالی اسـتفاده‌شـد. درنهایت مدل CLBS3 بر روی مجموعه داده UNSW-NB15 و TONIOT مورد آزمایش قرارگرفت.

نتایج آزمایش‌های مختلف نشـان‌می‌دهد که مدل پیشـنهادی در معیارهای نرخ تشـخیص بالا، دقت بالا و نرخ مثبت کاذب به ۹۸/۴۶٪، ۹۷/۹۹٪ و ۰,۰۰۳۲ در UNSW-NB15و ۹۸/۲٪ و ۹۸٪ و ۰/۰۰۲۸ در TONOT پیداکرده‌است که عملکرد بالای آن در تشـخیص حملات را اثبات‌می‌کند. طراحی و پیاده‌سـازی مدل برای حالت چندکلاسـه، همچنین اسـتفاده از تکنیک‌های بهینه‌سـازی هایپر پارامترها مدل پیشنهادی CLBS3 به کارهای آینده واگذارمی‌شود.

نتایج بیانگر آن اسـت که مدل پیشـنهادی برای کاربردهای IoT از نظر مصـرف منابع پردازشـی و حافظه بهینه طراحی‌شـده و می‌تواند در محیط‌هایی با محدودیت‌های سـخت‌افزاری اجراشـود. این مدل ضـمن بهره‌وری بـالا، بـار پردازشـی کمی بر CPU دارد، حـافظـه زیـادی اشـغال‌نمی‌کند و زمان اجرای کوتاهی دارد که همگی ویژگی های مناسبی برای دستگاه‌های IoT به شمار می‌آیند.

مراجع

[۱] P. Mishra, V. Varadharajan, U. Tupakula, and E. S. Pilli, "A detailed investigation and analysis of using machine learning techniques for intrusion detection," *IEEE communications surveys & tutorials*, vol. 21, pp. 686-728, 2018.

[۲] L. O. Anyanwu, J. Keengwe, and G. A. Arome, "Scalable intrusion detection with recurrent neural networks," in *2010 Seventh International Conference on Information Technology: New Generations*, 2010, pp. 919-923.

[۳] Z. X. Yang, X. L. Qin, W. R. Li, and Y. J. Yang, "A DDoS detection approach based on CNN in cloud computing," *Applied Mechanics and Materials*, vol. 513, pp. 579-584, 2014.

[۴] R. C. Staudemeyer, "Applying long short-term memory recurrent neural networks to intrusion detection," *South African Computer Journal*, vol. 56, pp. 136-154, 2015.

[۵] H. K. Maragheh, F. S. Gharehchopogh, K. Majidzadeh, and A. B. Sangar, "A Hybrid Model Based on Convolutional Neural Network and Long Short-Term Memory for Multi-label Text Classification," *Neural Processing Letters*, vol. 56, p. 42, 2024/02/16 2024.

[۶] H. Ebrahimi, K. Majidzadeh, and F. Soleimanian Gharehchopogh, "Integration of deep learning model and feature selection for multi-label classification," *International Journal of Nonlinear Analysis and Applications*, vol. 13, pp..۲۰۲۲ ,۲۸۸۳-۲۸۷۱

[۷] H. Wang, J. Gu, and S. Wang, "An effective intrusion detection framework based on SVM with feature augmentation," *Knowledge-Based Systems*, vol. 136, pp. 130-139, 2017.

[۸] M. Usha and P. Kavitha, "Anomaly based intrusion detection for 802.11 networks with optimal features using SVM classifier," *Wireless Networks*, vol. 23, pp. 2431-2446, 2017.

[۹] W. Meng, W. Li, and L. F. Kwok, "Design of intelligent KNN-based alarm filter using knowledge-based alert verification in intrusion detection," *Security and Communication Networks*, vol. 8, pp. 3883-3895, 2015.

[۱۰] G. Serpen and E. Aghaei, "Host-based misuse intrusion detection using PCA feature extraction and kNN classification algorithms," *Intelligent Data Analysis*, vol. 22, pp. 1101-1114,.۲۰۱۸

[۱۱] Y. Y. Aung and M. M. Min, "An analysis of K-means algorithm based network intrusion detection system," *Advances in Science, Technology and Engineering Systems Journal*, vol. 3, pp. 496-501, 2018.

[۱۲] N. Farnaaz and M. Jabbar, "Random forest modeling for network intrusion detection system," *Procedia Computer Science*, vol. 89, pp. 213-217, 2016.

[۱۳] K. Peng, V. C. Leung, L. Zheng, S. Wang, C. Huang, and T. Lin, "Intrusion detection system based on decision tree over big data in fog environment, " *Wireless Communications and Mobile Computing*, vol. 2018, p. 4680867, 2018.

[۱۴] E. Besharati, M. Naderan, and E. Namjoo, "LR-HIDS: logistic regression host-based intrusion detection system for cloud environments," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, pp. 3669-3692, 2019.

[۱۵] S. Teng, N. Wu, H. Zhu, L. Teng, and W. Zhang, "SVM-DT-based adaptive and collaborative intrusion detection," *IEEE/CAA Journal of Automatica Sinica*, vol. 5, pp. 108-118, 2017.

[۱۶] P. Tao, Z. Sun, and Z .Sun, "An improved intrusion detection algorithm based on GA and SVM," *Ieee Access*, vol. 6, pp. 13624-13631, 2018.

[۱۷] A. Khraisat, I. Gondal, P. Vamplew, J. Kamruzzaman, and A. Alazab, "A novel ensemble of hybrid intrusion detection system for detecting internet of things attacks," *Electronics*, vol. 8, p. 1210, 2019.

[۱۸] A. Shenfield, D. Day, and A. Ayesh, "Intelligent intrusion detection systems using artificial neural networks," *Ict Express*, vol. 4, pp. 95-99, 2018.

[۱۹] M. M. Baig, M. M. Awais, and E.-S. M. El-Alfy, "A multiclass cascade of artificial neural network for network intrusion detection," *Journal of Intelligent & Fuzzy Systems*, vol. 32, pp. 2875-2883, 2017.

[۲۰] I. Sumaiya Thaseen, J. Saira Banu, K. Lavanya, M. Rukunuddin Ghalib, and K. Abhishek, "An integrated intrusion detection system using correlation-based attribute selection and artificial neural network," *Transactions on Emerging Telecommunications Technologies*, vol. 32, p. e4014, 2021.

[۲۱] J. Kim, J. Kim, H. Kim, M. Shim, and E. Choi, "CNN-based network intrusion detection against denial-of-service attacks," *Electronics*, vol. 9, p. 916, 2020.

[۲۲] M. T. Nguyen and K. Kim, "Genetic convolutional neural network for intrusion detection systems," *Future Generation Computer Systems*, vol. 113, pp. 418-427, 2020.

[۲۳] S. A. Althubiti, E. M. Jones, and K. Roy, "LSTM for anomaly-based network intrusion detection," in *2018 28th International telecommunication networks and applications conference (ITNAC)*, 2018, pp. 1-3.

[۲۴] A. Chawla, P. Jacob, B. Lee, and S. Fallon, "Bidirectional LSTM autoencoder for sequence based anomaly detection in cyber security," *International Journal of Simulation–Systems, Science & Technology*, vol. 20, pp. 1-6, 2019.

[۲۵] M. Amar and B. E. Ouahidi, "Weighted LSTM for intrusion detection and data mining to prevent attacks," *International Journal of Data Mining, Modelling and Management*, vol. 12, pp. 308-329, 2020.

[۲۶] S. N. M. Khosroshahi, S. N. Razavi, A. Babazadeh Sangar, and K. Majidzadeh, "Offline Identification of the Author using Heterogeneous Data based on Deep Learning," *Computational Intelligence in Electrical Engineering*, vol. 13, pp. 115-134, 2022.

[۲۷] M. A. Khan, M. R. Karim, and Y. Kim, "A scalable and hybrid intrusion detection system based on the convolutional-LSTM network," *Symmetry*, vol. 11, p. 583, 2019.

[۲۸] W. Wang, Y. Sheng, J. Wang, X. Zeng, X. Ye, Y. Huang, *et al.*, "HAST-IDS: Learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection," *IEEE access*, vol. 6, pp. 1792-1806, 2017.

[۲۹] J. Zhang, Y. Ling, X. Fu, X. Yang, G. Xiong, and R. Zhang, "Model of the intrusion detection system based on the integration of spatial-temporal features," *Computers & Security*, vol. 89, p. 101681, 2020.

[۳۰] J. Canedo and A. Skjellum, "Using machine learning to secure IoT systems," in *2016 14th annual conference on privacy, security and trust (PST)*, 2016, pp. 219-222.

[۳۱] M. K. Putchala, "Deep learning approach for intrusion detection system (ids) in the internet of things (iot) network using gated recurrent neural networks (gru)," Wright State University, 2017.

[۳۲] M. Mohammadi, A. Al-Fuqaha, S. Sorour, and M. Guizani, "Deep learning for IoT big data and streaming analytics: A survey," *IEEE Communications Surveys & Tutorials*, vol. 20, pp. 2923-2960, 2018.

[۳۳] N. Moustafa, B. Turnbull, and K.-K. R. Choo, "An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things," *IEEE Internet of Things Journal*, vol. 6, pp. 4815-4830, 2018.

[۳۴] B. Roy and H. Cheung, "A deep learning approach for intrusion detection in internet of things using bi-directional long short-term memory recurrent neural network," in *2018 28th international telecommunication networks and applications conference (ITNAC)*, 2018, pp. 1-6.

[۳۵] T. S. Fatayer and M. N. Azara, "IoT secure communication using ANN classification algorithms," in *2019 International Conference on Promising Electronic Technologies (ICPET)*, 2019, pp. ۱۴۲-۱۴۶.

[۳۶] Y. N. Soe, Y. Feng, P. I. Santosa, R. Hartanto, and K. Sakurai, "A sequential scheme for detecting cyber attacks in IoT environment," in *2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*, 2019, pp. 238-244.

[۳۷] Z. Ma, J. Li, Y. Song, X. Wu, and C. Chen, "Network intrusion detection method based on FCWGAN and BiLSTM," *Computational Intelligence and Neuroscience*, vol. 2022, p. 6591140, 2022.

[۳۸] M. Saharkhizan, A. Azmoodeh, A. Dehghantanha, K.-K. R. Choo, and R. M. Parizi, "An ensemble of deep recurrent neural networks for detecting IoT cyber attacks using network traffic," *IEEE Internet of Things Journal*, vol. 7, pp. 8852-8859, 2020.

[۳۹] R. A. Ramadan and K. Yadav, "A novel hybrid intrusion detection system (IDS) for the detection of internet of things (IoT) network attacks," *Annals of Emerging Technologies in Computing (AETiC), Print ISSN*, pp. 2516-0281, 2020.

[۴۰] G. D. L. T. Parra, P. Rad, K.-K. R. Choo, and N. Beebe, "Detecting Internet of Things attacks using distributed deep learning," *Journal of Network and Computer Applications*, vol. 163, p.۲۰۲۰ ,۱۰۲۶۶۲ .

[۴۱] I. Ullah and Q. H. Mahmoud, "Design and development of a deep learning-based model for anomaly detection in IoT networks," *IEEE Access*, vol. 9, pp. 103906-103926, 2021.

[۴۲] M. Roknaldini and E. Noroozi, "Presenting A Hybrid Method of Deep Neural Networks to Prevent Intrusion in Computer Networks," *Intelligent Multimedia Processing and Communication Systems (IMPCS)*, vol. 4, pp. 57-65, 2023.

[۴۳] M. Nazarpour, N. Nezafati, and S. Shokouhyar, "Using the Modified Colonial Competition Algorithm to Increase the Speed and Accuracy of the Intelligent Intrusion Detection System," *Intelligent Multimedia Processing and Communication Systems (IMPCS)*, vol. 4, pp. 1-10, 2023.

[۴۴] A. Ghaffari and R. Hossinnezhad, "Intrusions detection system in the cloud computing using heterogeneity detection technique," *Intelligent Multimedia Processing and Communication Systems (IMPCS)*, vol. 3, pp. 37-46, 2022.

[۴۵] P. R. Kanna and P. Santhi, "Unified deep learning approach for efficient intrusion detection system using integrated spatial–temporal features," *Knowledge-Based Systems*, vol. 226, p. 107132, 2021.

[۴۶] P. Kumar, G. P. Gupta, and R. Tripathi, "An ensemble learning and fog-cloud architecture-driven cyber-attack detection framework for IoMT networks," *Computer Communications*, vol. 166, pp. 110-124, 2021.

[۴۷] S. Latif, Z. e Huma, S. S. Jamal, F. Ahmed, J. Ahmad, A. Zahid, *et al.*, "Intrusion detection framework for the internet of things using a dense random neural network," *IEEE Transactions on Industrial Informatics*, vol. 18, pp. 6435-6444, 2021.

[۴۸] M. Sarhan, S. Layeghy, N. Moustafa, M. Gallagher, and M. Portmann, "Feature extraction for machine learning-based intrusion detection in IoT networks," *Digital Communications and Networks*, 2022.

[۴۹] R. A. Elsayed, R. A. Hamada, M. I. Abdalla, and S. A. Elsaid, "Securing IoT and SDN systems using deep-learning based automatic intrusion detection," *Ain Shams Engineering Journal*, vol. 14, p. 102211, 2023.

[۵۰] R. Anushiya and V. Lavanya, "A new deep-learning with swarm based feature selection for intelligent intrusion detection for the Internet of things," *Measurement: Sensors*, vol. 26, p. 100700, 2023.

[۵۱] R. Zhao, G. Gui, Z. Xue, J. Yin, T. Ohtsuki, B. Adebisi, *et al.*, "A novel intrusion detection method based on lightweight neural network for internet of things," *IEEE Internet of Things Journal*, vol. 9, pp. 9960-9972, 2021.

[۵۲] S. M. Kasongo, "A deep learning technique for intrusion detection system using a Recurrent Neural Networks based framework," *Computer Communications*, vol. 199, pp. 113-125, 2023.

پی‌نوشت

[1] Convolutional neural network and Long Short Term Memory Branch Splitter with three dimensions' data(CLBS3)

[2] False Positive Rate (FPR)

[3] False Negative Rate (FNR)

[4] True Positive(TP)

[5] True Negative(TN)

[6] False Negative(FN)