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

Summer 2024, 5 (2), 177-192
DOR:

Received: 16 May 2024
Accepted: 30 June 2024

مقاله پژوهشی

A Hybrid Model for Intrusion Detection Using Big Data Analytics and Deep Learning Techniques in Distributed Networks

Shahin SamieAdel*

Ph.D. Student, Department of Management, Najafabad Branch, Islamic Azad University, Najafabad, Iran. **Corresponding Author*, shahinsamieadel@gmail.com

Abstract

This research presents a new hybrid model for intrusion detection in distributed environments, based on big data analytics and deep learning. The primary goal of the study is to develop a system with high accuracy and efficiency for rapid identification and response to cyberattacks. The methodology includes preprocessing network traffic data and utilizing convolutional and recurrent neural networks to detect complex attack patterns. The proposed model, evaluated on the KDD Cup 99 dataset, demonstrated remarkable performance with 98% accuracy, a 95% detection rate, and a false positive rate of less than 2%. The model's ability for continuous learning and automatic updates enables it to adapt to emerging threats. Additionally, by incorporating advanced encryption methods and federated learning techniques, sensitive data security is ensured. This study can serve as a foundation for future research in cybersecurity and the development of intelligent intrusion detection systems.

Introduction : In the current digital age, Cybersecurity has become one of the main concerns of various organizations and institutions. With the increasing complexity and number of cyber threats, the necessity of employing advanced systems to identify and counter these threats becomes more apparent. Intrusion Detection refers to the process of detecting unauthorized attempts to access computer systems and networks. These systems attempt to identify unusual patterns in network traffic and prevent attacks. Utilizing Deep Learning techniques and Big Data analysis enables the design of more advanced intrusion detection systems capable of detecting more complex attacks[1,2]. Distributed and Cloud networks are among the emerging and widely used technologies that are rapidly expanding due to their superior capabilities in data processing and storage. These networks allow companies and institutions to offer their services with greater efficiency and flexibility. However, this rapid and extensive development of distributed and cloud networks also brings new security challenges. Among these challenges is the need for efficient and advanced intrusion detection systems to identify and respond to cyberattacks [3]. One effective approach to enhance the accuracy and efficiency of intrusion detection systems in distributed environments is the use of Hybrid Models of deep learning and Big Data analysis. These models can identify cyberattacks with greater accuracy by analyzing complex and sequential patterns in network data. Recent studies have shown that using deep learning models can significantly improve the performance of intrusion detection systems [4]. The present research focuses on the design and implementation of a hybrid intrusion detection system in distributed environments. The proposed system, using big data analysis and deep learning, has the capability to identify and respond to cyberattacks in real-time. The main objective of this paper is to present a model with high accuracy and efficiency to meet the security needs of distributed and cloud networks, and in addition to examining the theoretical foundations and literature review, it evaluates the proposed model with real data. The results of this research can help improve intrusion detection systems. In this study, special attention has been paid to the issue of system scalability. Given the vast volume of data generated in modern networks, the capability for parallel and distributed processing has been incorporated into the system design. This feature enables the use of the system in large and complex environments. Additionally, one of the innovations of this research is the integration of reinforcement learning techniques with deep learning models. This approach allows the system to continuously learn from its experiences and improve its performance. This capability is particularly useful in dealing with emerging and unknown threats. Furthermore, the present research examines the ethical and legal aspects of using artificial intelligence in security systems. This topic is of particular importance given the increasing concerns about privacy and ethical use of advanced technologies.

Method: This research generally consists of two parts: review and exploratory. In the review section, a comprehensive examination of research literature in the field of intrusion detection systems and deep learning models was conducted, and scientific articles from the last five years from reputable databases such as IEEE, Springer, and Elsevier were reviewed. The aim of this section was to identify research related to the optimization of intrusion detection models and analyze gaps in past studies. Additionally, Google Scholar and PubMed search tools were used to find new articles related to Cybersecurity and deep learning. In the exploratory section, data were collected from two main sources including scientific articles and real datasets related to cyberattacks such as KDD Cup 99 and NSL-KDD. These sets are well-suited for evaluating and analyzing the proposed models and contain information on

various cyberattacks. The data preprocessing process included cleaning and normalizing them to remove duplicate and invalid data and extract key features related to cyberattacks. In the implementation of deep learning models, convolutional neural networks and recurrent neural networks were used. The parameters of these models were adjusted based on the number of layers, learning rate, and number of neurons, and optimization techniques such as Adam and Dropout were employed to improve performance and prevent overfitting. Model evaluation was performed using metrics such as Accuracy, detection rate, and false positive rate. To improve the accuracy and efficiency of the models, methods such as Cross-Validation and transfer learning were employed to enable knowledge transfer between different models and achieve optimal results. Python was used as the main language for model implementation and data analysis, and tools such as TensorFlow and Keras were used as deep learning frameworks. Additionally, R was used for statistical analysis and initial analyses. This multidimensional approach and the use of various tools have provided the possibility of presenting an efficient and accurate model for detecting and countering cyberattacks in distributed environments.

Results: The main objective of this research was to develop and evaluate a hybrid intrusion detection model using deep learning techniques and big data analysis in distributed environments. The proposed model combined convolutional neural networks (CNN) for spatial feature extraction and recurrent neural networks (RNN) for temporal pattern analysis from network traffic data. The evaluation was conducted using the KDD Cup 99 dataset, which includes diverse categories of attacks, such as DoS, R2L, U2R, and Probe attacks. The CNN-RNN hybrid model demonstrated excellent performance, achieving a detection accuracy of 98% and a false positive rate of less than 2%. The model successfully identified both known and complex unknown attack patterns, leveraging big data analysis to process large volumes of network traffic efficiently.

Key Results:

- Detection accuracy of 98% in identifying cyber-attacks.
- False positive rate reduced to less than 2%, crucial for minimizing false alarms.
- Efficient detection of complex and sequential attack patterns like DoS and R2L attacks.
- These findings confirm the effectiveness of the proposed model in improving the accuracy and efficiency of intrusion detection systems.

Discussion: The results of this study indicate that combining deep learning models with big data analysis significantly improves the performance of intrusion detection systems, particularly in distributed environments. The CNN-RNN hybrid model efficiently captured both spatial and temporal patterns in network traffic, enabling it to detect complex cyber-attacks with high accuracy. The ability to reduce the false positive rate to less than 2% is a critical achievement, as it helps minimize unnecessary alerts and enhances the reliability of security systems. Furthermore, the use of the KDD Cup 99 dataset, which covers a wide range of cyber-attack scenarios, validated the robustness of the model. The high detection rate of over 95% for different attack types, including DoS and R2L, demonstrates the model’s capability to identify both traditional and evolving threats in real-time environments. The hybrid model's combination of deep learning and big data analysis allows for scalable solutions in advanced cybersecurity frameworks, making it suitable for real-world applications where rapid and accurate threat detection is essential. In conclusion, the hybrid CNN-RNN model shows great promise in enhancing the efficiency of intrusion detection systems, providing a reliable tool for defending against emerging and complex cyber threats.

Keywords: Intrusion Detection, Cybersecurity, Deep Learning, Big Data, Distributed Environments.

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

مدل ترکیبی برای کشف نفوذ با استفاده از تحلیل داده‌های بزرگ و تکنیک‌های یادگیری عمیق در شبکه‌های توزیع‌شده

تاریخ دریافت: ۲۷/ ۰۲/ ۱۴۰۳
تاریخ پذیرش: ۱۰/ ۰۴/ ۱۴۰۳

دورهٔ پنجم، تابستان ۱۴۰۳
شمارهٔ دوم، صص: ۱۷۷-۱۹۲

شاهین سمیع عادل *

* دانشجوی دکتری، گروه مدیریت، واحد نجف آباد، دانشگاه آزاد اسلامی، نجف آباد، ایران. (نویسنده مسئول) shahinsamieadel@gmail.com

چکیده: این پژوهش مدل ترکیبی جدیدی برای کشف نفوذ در محیط‌های توزیع‌شده ارائه‌می‌دهد که بر پایه تحلیل داده‌های بزرگ و یادگیری عمیق عمل‌می‌کند. هدف اصلی این مطالعه، توسعه سیستمی با دقت و کارایی بالا برای شناسایی و پاسخ سریع به حملات سایبری است. روش شناسی شامل پیش‌پردازش داده‌های ترافیک شبکه و استفاده از شبکه‌های عصبی پیچشی و بازگشتی جهت شناسایی الگوهای پیچیده حملات است. ارزیابی مدل پیشنهادی با داده‌های KDD Cup 99 نشان داد که سیستم با دقت ۹۸٪، نرخ کشف ۹۵٪ و نرخ مثبت کاذب کمتر از ۲٪، عملکرد قابل‌توجهی دارد. این مدل با قابلیت یادگیری مداوم و به‌روزرسانی خودکار می‌تواند در برابر تهدیدات نوظهور تطبیق‌یابد. همچنین، با ادغام روش‌های رمزنگاری پیشرفته و یادگیری فدرال، امنیت داده‌های حساس تضمین شده‌است. این مطالعه می‌تواند به عنوان مبنایی برای تحقیقات آینده در حوزه امنیت سایبری و توسعه سیستم‌های کشف نفوذ هوشمند مورد استفاده قرارگیرد.

واژه‌های کلیدی: کشف نفوذ، امنیت سایبری، یادگیری عمیق، داده‌های بزرگ، محیط‌های توزیع‌شده.

۱. مقدمه

در عصر دیجیتال کنونی، امنیت سایبری[۱] به یکی از دغدغه‌های اصلی سازمان‌ها و نهادهای مختلف تبدیل‌شده‌است. با افزایش پیچیدگی و تعدد تهدیدات سایبری، ضرورت به‌کارگیری سیستم‌های پیشرفته برای شناسایی و مقابله با این تهدیدات بیش از پیش نمایان‌می‌شود. کشف نفوذ[۲] به فرآیند تشخیص تلاش‌های غیرمجاز برای دستیابی به سیستم‌های رایانه‌ای و شبکه‌ها اطلاق‌می‌شود. این سیستم‌ها تلاش‌دارند الگوهای غیرمعمول را در ترافیک شبکه شناسایی‌کرده و از بروز حملات پیشگیری نمایند. بهره‌گیری‌از تکنیک‌های یادگیری عمیق[۳] و تحلیل داده‌های بزرگ، امکان طراحی سیستم‌های کشف نفوذ پیشرفته‌تری را فراهم‌می‌آورد که قابلیت تشخیص حملات پیچیده‌تر را دارا هستند [۱،۲]. شبکه‌های توزیع‌شده[۴] و ابری[۵] از جمله فناوری‌های نوظهور و پرکاربردی محسوب‌می‌شوند که به‌واسطه قابلیت‌های برتر خود در پردازش و ذخیره‌سازی داده‌ها، به‌سرعت در حال گسترش می‌باشند. این شبکه‌ها به شرکت‌ها و موسسات امکان‌می‌دهند خدمات خود را با بهره‌وری و انعطاف‌پذیری بیشتری عرضه‌کنند. با این وجود، این توسعه سریع و گسترده شبکه‌های توزیع‌شده و ابری، چالش‌های امنیتی نوینی را نیز به همراه دارد. از جمله این چالش‌ها، نیاز به سیستم‌های کشف نفوذ کارآمد و پیشرفته برای شناسایی و واکنش به حملات سایبری است [۳]. یکی از رویکردهای موثر برای ارتقای دقت و کارایی سیستم‌های کشف نفوذ در محیط‌های توزیع‌شده، استفاده از مدل‌های ترکیبی[۶] یادگیری عمیق و تحلیل داده‌های بزرگ[۷] است. این مدل‌ها قادرند با تحلیل الگوهای پیچیده و متوالی در داده‌های شبکه، حملات سایبری را با صحت بیشتری شناسایی کنند. مطالعات اخیر نشان‌داده‌اند که استفاده از مدل‌های یادگیری عمیق می‌تواند به‌طور قابل‌ملاحظه‌ای عملکرد سیستم‌های کشف نفوذ را بهبود بخشد [۴].

پژوهش حاضر به طراحی و پیاده‌سازی یک سیستم کشف نفوذ ترکیبی در محیط‌های توزیع‌شده پرداخته‌است که با استفاده از تحلیل داده‌های بزرگ و یادگیری عمیق، قابلیت شناسایی و پاسخ به حملات سایبری در زمان واقعی را فراهم‌می‌کند. هدف اصلی این تحقیق ارائه مدلی با دقت و کارایی بالا برای تأمین نیازهای امنیتی شبکه‌های توزیع‌شده و ابری است. علاوه بر بررسی مبانی نظری و مرور ادبیات، ارزیابی مدل پیشنهادی بر اساس داده‌های واقعی صورت گرفته‌است تا قابلیت آن در شرایط واقعی سنجیده‌شود.

یکی از نوآوری‌های کلیدی این پژوهش، طراحی مدلی ترکیبی است که از شبکه‌های عصبی پیچشی و بازگشتی برای تحلیل هم‌زمان ویژگی‌های مکانی و زمانی ترافیک شبکه استفاده‌می‌کند. این ویژگی، دقت مدل را در شناسایی حملات پیچیده افزایش داده و به کاهش نرخ هشدارهای مثبت کاذب کمک‌کرده‌است. همچنین، برای ارتقای مقیاس‌پذیری سیستم و بهینه‌سازی عملکرد در پردازش حجم عظیم داده‌های شبکه‌های توزیع‌شده، از پردازش موازی و توزیع‌شده استفاده‌شده‌است که این سیستم را برای محیط‌های بزرگ و پیچیده مناسب می‌سازد. نوآوری دیگر این پژوهش، ادغام یادگیری تقویتی با مدل‌های یادگیری عمیق است که به سیستم اجازه می‌دهد به‌صورت مداوم از تجربیات خود یادبگیرد و در برابر تهدیدات سایبری نوظهور و ناشناخته عملکرد بهتری داشته‌باشد. به دلیل نگرانی‌های مربوط به حریم خصوصی و استفاده اخلاقی از هوش مصنوعی، این پژوهش به بررسی جنبه‌های اخلاقی و حقوقی بهره‌گیری از فناوری‌های پیشرفته در سیستم‌های امنیتی نیز پرداخته‌است، تا راهکارهای مناسبی برای استفاده مسئولانه از این فناوری ارائه‌دهد.

در حالی‌که مدل‌های کشف نفوذ فعلی به پیشرفت‌های قابل‌توجهی در دقت دست‌یافته‌اند، چالش‌های کلیدی همچنان پابرجاست. این چالش‌ها شامل نرخ بالای هشدارهای مثبت کاذب، عدم توانایی در شناسایی حملات جدید و پیچیده، و نیاز به منابع محاسباتی قابل‌توجه برای پردازش بلادرنگ است. برای مثال، مدل‌های مبتنی بر یادگیری عمیق اغلب در تحلیل داده‌های نامتوازن با دشواری مواجه هستند. تحقیق حاضر تلاش‌می‌کند با ترکیب روش‌های تحلیل داده‌های بزرگ و مدل‌های یادگیری عمیق ترکیبی (CNN-RNN)، به بهبود عملکرد و کارایی در این زمینه دست‌یابد.

۲. بیان مسئله و اهداف پژوهش

گسترش سریع فناوری اطلاعات و ارتباطات، منجر به افزایش حجم داده‌های تولیدی شده و این رشد سریع چالش‌های جدیدی را برای سیستم‌های کشف نفوذ به وجود آورده‌است. سیستم‌های سنتی در پردازش و تحلیل حجم بالای داده‌های شبکه با محدودیت‌هایی مواجه هستند و در شناسایی حملات پیچیده و سایبری ناتوان عمل‌می‌کنند. از این رو، به‌کارگیری مدل‌های پیشرفته‌تر برای شناسایی و مقابله با این حملات ضروری است. هدف این پژوهش طراحی و توسعه مدل پیشرفته کشف نفوذ است که با استفاده از تکنیک‌های یادگیری عمیق و تحلیل داده‌های بزرگ با دقت و کارآیی بالا در محیط‌های توزیع‌شده عمل‌می‌کند. این مدل با بهره‌گیری از الگوریتم‌های پیشرفته هوش مصنوعی و پردازش توزیع‌شده، توانایی تحلیل حجم عظیمی از داده‌ها را در زمان واقعی فراهم‌می‌آورد و به شناسایی و مقابله با حملات سایبری پیچیده کمک‌می‌کند. از جمله نوآوری‌های کلیدی این مدل، تلفیق یادگیری انتقالی با یادگیری عمیق است، که به سیستم امکان‌می‌دهد دانش به‌دست‌آمده از یک حوزه را به سایر حوزه‌ها منتقل‌کند و در نتیجه کارایی و انعطاف‌پذیری آن افزایش‌یابد.

سیستم‌های کشف نفوذ[۸] ابزاری برای شناسایی فعالیت‌های مشکوک در شبکه‌ها هستند و به دو دسته کلی تقسیم‌می‌شوند: سیستم‌های کشف نفوذ مبتنی بر امضا[۹] که با تطابق الگوها به شناسایی حملات می‌پردازند و سیستم‌های کشف نفوذ مبتنی بر رفتار[۱۰] که فعالیت‌های غیرعادی شبکه را شناسایی‌می‌کنند [۵]. پژوهش‌ها نشان می‌دهند که روش‌های مبتنی بر یادگیری ماشین، قابلیت بهبود دقت و کارایی این سیستم‌ها را دارند [۶]. با این حال افزایش حجم داده‌ها و پیچیدگی حملات، نیاز به تکنیک‌های پیشرفته‌تر مانند یادگیری عمیق را ضروری کرده‌است [۷].

یادگیری عمیق به‌عنوان شاخه‌ای از یادگیری ماشین، با استفاده از شبکه‌های عصبی چند لایه، الگوهای پیچیده‌تری را از داده‌های بزرگ استخراج‌می‌کند. این روش‌ها در مسائل مختلفی مانند تشخیص تصویر و شناسایی الگوهای غیرعادی در شبکه‌ها به کار گرفته‌می‌شوند. مدل‌های یادگیری عمیق از جمله شبکه‌های عصبی پیچشی و بازگشتی، قابلیت شناسایی الگوهای متوالی و پیچیده را در داده‌های شبکه دارند و می‌توانند به‌طور مؤثر برای کشف نفوذ استفاده شوند [۸]. تحلیل داده‌های بزرگ به استخراج الگوهای معنادار از مجموعه‌های داده حجیم و پیچیده اشاره دارد. با توجه به افزایش حجم داده‌های تولید شده در محیط‌های توزیع‌شده، اهمیت این روش‌ها در کشف نفوذ افزایش یافته است [۹]. ترکیب یادگیری عمیق با تحلیل داده‌های بزرگ به شکل مدل‌های ترکیبی مانند شبکه‌های عصبی پیچشی-بازگشتی، به شناسایی الگوهای پیچیده و حملات سایبری کمک‌می‌کند [۴].

۳. اهمیت موضوع

محیط‌های توزیع‌شده و ابری نیز به دلیل قابلیت‌های پردازش و ذخیره‌سازی گسترده خود، رشد سریعی داشته‌اند، اما این توسعه چالش‌های امنیتی جدیدی را نیز به همراه دارد [۴]. به همین دلیل رو به‌روز نیاز به سیستم‌های کشف نفوذ توزیع‌شده که قادر به شناسایی حملات سایبری در زمان واقعی باشند، بیشتر می‌شود. در ارزیابی سیستم‌های کشف نفوذ، معیارهایی مانند دقت[۱۱]، نرخ کشف[۱۲] و نرخ مثبت کاذب[۱۳] استفاده‌می‌شود. دقت، توانایی سیستم در شناسایی حملات واقعی را نشان‌می‌دهد، در حالی که نرخ مثبت کاذب به هشدارهای نادرستی اشاره‌دارد که سیستم به اشتباه به عنوان حمله تشخیص‌می‌دهد. نرخ مثبت کاذب بالا به کاهش اعتماد به سیستم و مصرف منابع بیشتر برای تحلیل هشدارهای نادرست منجرمی‌شود. مطالعات نشان می‌دهند مدل‌های ترکیبی مبتنی بر یادگیری عمیق و تحلیل داده‌های بزرگ، عملکرد بهتری نسبت به روش‌های سنتی دارند [۷،۹]. یکی از نقاط ضعف مطالعات پیشین، عدم‌وجود چارچوبی استاندارد برای ارزیابی و مقایسه مدل‌های مختلف است. مطالعات متعددی از مجموعه داده‌ها و روش‌های ارزیابی متفاوتی استفاده‌کرده‌اند که مقایسه مستقیم نتایج را دشوار می‌سازد. به همین منظور در مطالعه دیگری [۱۰] با مرور سیستماتیک و متاآنالیز به بررسی و تحلیل داده‌ها پرداخته و شکاف‌های پژوهش‌های گذشته مشخص‌شد. این بررسی نشان‌می‌دهد که مدل‌های ترکیبی با یادگیری عمیق به بهبود قابل توجه در سیستم‌های کشف نفوذ منجر می‌شوند.

۴. روش‌شناسی پژوهش

این پژوهش به‌طور کلی از دو بخش مروری و اکتشافی تشکیل شده است. در بخش مروری، بررسی مقالات پنج سال اخیر در زمینه سیستم‌های کشف نفوذ و یادگیری عمیق از پایگاه‌های IEEE، Springer و Elsevier انجام و از Google Scholar و PubMed انجام‌شد. هدف از این جستجو، شناسایی پژوهش‌های مرتبط با بهینه‌سازی مدل‌های کشف نفوذ و تحلیل شکاف‌های موجود بود. در بخش اکتشافی، داده‌ها از مقالات علمی و مجموعه داده‌های واقعی مانند KDD Cup 99 و NSL-KDD جمع‌آوری‌شدند که برای ارزیابی و تحلیل مدل‌های پیشنهادی مناسب بوده و حاوی اطلاعات حملات مختلف سایبری هستند. در انتخاب مجموعه داده KDD Cup 99، دلایل متعددی مدنظر قرارگرفته‌است. مجموعه داده KDD Cup 99 به دلیل تنوع حملات سایبری و ساختار مناسب، به عنوان استانداردی معتبر در ارزیابی مدل‌های کشف نفوذ انتخاب‌شد. این مجموعه که شامل داده‌های مرتبط با حملاتی مانند DoS، R2L، U2R و Probe است، به مدل‌های یادگیری عمیق کمک‌می‌کند تا الگوهای پیچیده حملات را بهتر شناسایی‌کنند و ارزیابی مدل را دقیق‌تر و قابل‌مقایسه‌تر سازد.

فرآیند پیش‌پردازش داده‌ها شامل پاک‌سازی و نرمال‌سازی برای حذف داده‌های تکراری و نامعتبر و استخراج ویژگی‌های کلیدی مرتبط با حملات سایبری بود. در بخش اجرای مدل‌های یادگیری عمیق، از شبکه‌های عصبی پیچشی و بازگشتی استفاده‌شد. پارامترهای این مدل‌ها بر اساس تعداد لایه‌ها، نرخ یادگیری و تعداد نرون‌ها تنظیم‌شده و تکنیک‌های بهینه‌ساز مانند Adam و Dropout برای بهبود عملکرد به کار گرفته‌شد. ارزیابی مدل‌ها با استفاده از معیارهایی همچون دقت، نرخ کشف و نرخ مثبت کاذب انجام‌شد و برای بهبود دقت، از روش‌های Cross-Validation و یادگیری انتقالی استفاده‌شد. پیاده‌سازی با Python و فریم‌ورک‌های TensorFlow و Keras، همراه با تحلیل‌های آماری در R انجام‌شد که این رویکرد چندبعدی، امکان ارائه مدلی کارآمد برای شناسایی حملات سایبری در محیط‌های توزیع‌شده را فراهم‌کرد.

انتخاب این مجموعه داده‌های KDD Cup 99 و NSL-KDD بر اساس سه ویژگی کلیدی انجام شده‌است:

۱. تنوع: این مجموعه داده‌ها شامل چهار نوع حمله اصلی (DoS، R2L، U2R، Probe) است که بازتابی از تهدیدات متنوع سایبری است.

۲. توزیع نامتوازن: در KDD Cup 99، حدود ۷۹٪ داده‌ها مربوط به حملات DoS هستند، در حالی‌که دسته‌های R2L و U2R کمتر از ۱٪ داده‌ها را شامل می‌شوند. این توزیع نشان‌دهنده یک چالش واقعی در داده‌های امنیتی است.

۳. معیار استاندارد: این مجموعه‌ها به‌عنوان داده‌های مرجع در ارزیابی سیستم‌های کشف نفوذ مورداستفاده قرارگرفتند و مقایسه عملکرد مدل‌ها را تسهیل می‌کنند.

برای مقابله با چالش توزیع نامتوازن، از تکنیک‌هایی مانند Oversampling و Weighted Loss استفاده شده‌است.

فرآیند پیش‌پردازش داده‌ها شامل پاک‌سازی (حذف داده‌های تکراری و غیرمعتبر)، نرمال‌سازی (مقیاس‌دهی ویژگی‌ها به محدوده استاندارد) و استخراج ویژگی‌های کلیدی نظیر مدت زمان ارتباط، نرخ ارسال بسته‌ها و تعداد بایت‌ها می‌باشد. مدل پیشنهادی شامل ترکیبی از شبکه‌های عصبی پیچشی (CNN) و بازگشتی (RNN) است. تنظیمات پارامترها شامل تعداد لایه‌های پیچشی (سه لایه با فیلترهای ۳۲، ۶۴ و ۱۲۸)،

نرخ یادگیری ۰,۰۰۱، و مکانیزم Dropout با نرخ ۰,۵ برای جلوگیری از بیش‌برازش بوده‌است. به‌منظور ارزیابی مدل، از Cross-Validation پنج‌برابر استفاده‌شد تا عملکرد آن در داده‌های مختلف سنجیده‌شود.

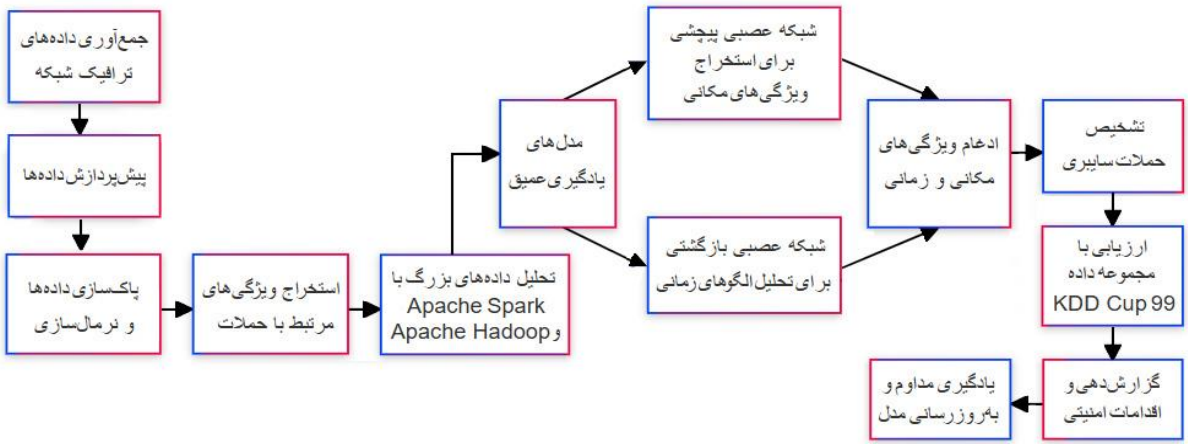
در این پژوهش پارامترهای کلیدی مدل‌های یادگیری عمیق به‌دقت تنظیم‌شده‌اند تا عملکرد بهینه حاصل‌شود. مثلاً تعداد لایه‌های شبکه‌های عصبی و نرخ یادگیری در مدل CNN-RNN برای دستیابی به تعادل مناسب میان دقت و سرعت تنظیم‌شده‌است. انتخاب پارامترهایی مانند تعداد نرون‌ها، نرخ ریزش[۱۴] و روش‌های بهینه‌سازی نظیر Adam ، به‌گونه‌ای صورت‌گرفته که از بیش‌برازش مدل جلوگیری‌شود و در عین حال بالاترین میزان دقت به‌دست‌آید .

۵. مدل‌های کشف نفوذ

در حوزه کشف نفوذ و امنیت سایبری، مدل‌های مبتنی بر یادگیری عمیق نقش مهمی در تشخیص و پیشگیری از حملات سایبری دارند. این مدل‌ها با بهره‌گیری از تکنیک‌های مختلف یادگیری عمیق، توانایی شناسایی و تحلیل الگوهای پیچیده را بهبود می‌بخشند و دقت بیشتری در تشخیص حملات ارائه‌می‌دهند، زیرا شبکه‌های عصبی کلاسیک برای سیستم تشخیص نفوذ هوشمند از دقت کافی برخوردار نیستند [۱۱]. در این میان انواع مختلفی از مدل‌های عصبی توسعه‌یافته که هرکدام برای شناسایی جنبه‌های خاصی از حملات سایبری مناسب هستند که در ادامه به بررسی چند مدل کلیدی در این زمینه می‌پردازیم.

۱,۵.مدل شبکه‌های عصبی بازگشتی[۱۵]

شبکه‌های عصبی بازگشتی نیز یکی دیگر از مدل‌های مهم در حوزه کشف نفوذ و به‌طور ویژه برای پردازش داده‌های ترتیبی و زمانی مناسب هستند. شبکه‌های عصبی بازگشتی با داشتن حافظه‌ای که می‌تواند اطلاعات قبلی را ذخیره‌کند، قادر به شناسایی الگوهای زمانی در داده‌های شبکه هستند. این قابلیت باعث‌می‌شود تا شبکه‌های عصبی بازگشتی بتوانند حملاتی که به مرور زمان شکل‌می‌گیرند، شناسایی‌کنند. از جمله کاربردهای موفق این مدل‌ها در کشف نفوذ می‌توان به شناسایی حملات منع سرویس توزیع‌شده[۱۶] و تحلیل ترافیک شبکه اشاره‌کرد [۲]. استفاده از مدل شبکه‌های عصبی بازگشتی در کشف نفوذ به دلیل توانایی آن در حفظ اطلاعات زمانی، مزایای زیادی دارد. بر اساس مطالعات جدید، این مدل‌ها در شناسایی حملات توزیع‌شده مانند DDoS بسیار موفق عمل‌کرده‌اند و توانسته‌اند تا ۹۸,۷٪ دقت در تشخیص حملات را نشان‌دهند [۱۲]. این مدل‌ها همچنین قابلیت بهینه‌سازی برای پردازش بلادرنگ را دارند.



شکل ۱- روش پژوهش

۲،۵.مدل شبکه‌های عصبی پیچشی[۱۷]

یکی از مدل‌های پرکاربرد در کشف نفوذ، شبکه‌های عصبی پیچشی است. این مدل‌ها در ابتدا برای پردازش تصاویر توسعه یافته‌بودند اما به دلیل توانایی بالایی که در استخراج ویژگی‌های محلی و شناسایی الگوهای پیچیده دارند، به‌سرعت در زمینه‌های دیگر نیز مورد استفاده قرارگرفتند. می‌توانند ویژگی‌های پیچیده و غیرخطی را از داده‌های ورودی به صورت خودکار استخراج‌کنند. شبکه‌های باور عمیق توانایی شناسایی دقیق حملات و حفاظت از ترافیک شبکه را دارند [۵]. شبکه‌های باور عمیق از ساختار چندلایه خود برای شناسایی ویژگی‌های پیچیده استفاده‌می‌کنند. بر اساس یک مطالعه اخیر، استفاده از مدل شبکه‌های باور عمیق در

جدول ۱- مقایسه مدل‌های کشف نفوذ [۱۷-۲۰]

معیارها	مدل شبکه‌های عصبی پیچشی	مدل شبکه‌های عصبی بازگشتی	مدل شبکه‌های باور عمیق	مدل شبکه‌های عصبی ترکیبی	مدل‌های مبتنی بر یادگیری انتقالی	مدل‌های ترکیبی و سیستم‌های توزیع‌شده
معماری	لایه‌های پیچشی و تجمعی برای استخراج ویژگی‌ها	شبکه‌های بازگشتی با حلقه‌ها برای داده‌های توالی	شبکه‌ها با لایه‌های پنهان چندگانه، استفاده از مدل‌های بولتزمن محدودشده	ترکیبی از مدل‌های عصبی مختلف مانند CNN و RNN	ترکیب مدل‌های پیش‌آموزش داده شده با داده‌های جدید	ترکیب مدل‌های مختلف در محیط‌های توزیع‌شده
قوت‌ها	کارآمد در تحلیل داده‌های مکانی، استخراج ویژگی‌های موثر	کارآمد در تحلیل داده‌های توالی و سری زمانی	کارآمد در یادگیری ویژگی‌های عمیق و مدل‌های پیچیده	بهره‌گیری از قوت‌های مدل‌های مختلف	بهبود دقت با استفاده از مدل‌های پیش‌آموزش داده شده	استفاده از منابع محاسباتی توزیع‌شده برای افزایش کارایی و مقیاس‌پذیری
ضعف‌ها	نیاز به داده‌های بزرگ، محاسبات فشرده	مشکل در آموزش توالی‌های بلند به دلیل مشکل شیب ناپدید	نیاز به تنظیم دقیق و محاسبات پیچیده	پیچیدگی بالا و نیاز به تنظیم دقیق	نیاز به داده‌های پیش‌آموزش داده شده مناسب	پیچیدگی بالا در مدیریت منابع و هماهنگی بین گره‌ها
مقیاس پذیری	مقیاس‌پذیری بالا با زیرساخت مناسب	مقیاس‌پذیری با روش‌های بهینه‌سازی آموزش	مقیاس‌پذیری محدود به منابع محاسباتی موجود	مقیاس‌پذیری بالا با ترکیب مدل‌های مختلف	مقیاس‌پذیری بالا با استفاده از مدل‌های پیش‌آموزش داده شده	مقیاس‌پذیری بسیار بالا در محیط‌های توزیع‌شده
پردازش بلادرنگ	امکان‌پذیر اما نیاز به بهینه‌سازی	مناسب برای پردازش بلادرنگ	محدودیت پردازش بلادرنگ به دلیل پیچیدگی مدل‌ها	مناسب پردازش بلادرنگ با بهینه‌سازی مناسب	امکان‌پذیر با استفاده از مدل‌های سبک‌تر	مناسب پردازش بلادرنگ در محیط‌های توزیع‌شده
دقت	دقت بالا در وظایف مکانی و تصویری	دقت بالا در وظایف توالی و سری زمانی	دقت بالا در وظایف پیچیده و داده‌های حجیم	دقت بالا با ترکیب قوت‌های مدل‌های مختلف	دقت بالا با استفاده ازداده‌های پیش‌آموزش داده شده مناسب	دقت‌بالا با ترکیب مدل‌های مختلف و توزیع‌شده
پیچیدگی	پیچیدگی بالا، نیاز به GPU	پیچیدگی‌متوسط تا بالا	پیچیدگی بالا، نیاز به تنظیم دقیق	پیچیدگی بسیار بالا به دلیل ترکیب مدل‌ها	پیچیدگی بالا در تنظیم و آموزش مجدد مدل‌ها	پیچیدگی بسیار بالا در مدیریت‌وهماهنگی‌بین گره‌ها
موارد استفاده	تشخیص تصویر، تشخیص ناهنجاری در تصاویر	پیش‌بینی سری‌های زمانی، مدل‌سازی زبان	تشخیص بیماری‌ها، پیش‌بینی‌وضعیت‌های پیچیده	تشخیص ناهنجاری‌ها، تحلیل داده‌های چندمنظوره	بهبود مدل‌های موجود، تشخیص ناهنجاری‌ها	تشخیص ناهنجاری‌ها در سیستم‌های توزیع‌شده، تحلیل داده‌های بزرگ

در سیستم‌های کشف نفوذ، شبکه‌های عصبی پیچشی می‌توانند الگوهای غیرعادی در ترافیک شبکه را با دقت بالا شناسایی‌کنند. این مدل‌ها از چندین لایه پیچشی و لایه‌های کاملاً متصل برای استخراج و ترکیب ویژگی‌ها استفاده‌می‌کنند. تحقیقات نشان‌داده‌اند که شبکه‌های عصبی پیچشی می‌توانند دقت بالایی در شناسایی حملات داشته و زمان پردازش را نیز بهبود بخشند [۱۳]. شبکه‌های عصبی پیچشی از قدرت بالایی در تشخیص الگوهای مکانی برخوردارند و توانسته‌اند در مسائلی چون تشخیص تصویر نیز به‌خوبی عمل‌کنند. تحقیقات اخیر نشان‌داده‌اند که استفاده از مدل شبکه‌های عصبی پیچشی در سیستم‌های کشف نفوذ می‌تواند تا ۹۹،۵٪ دقت را در تشخیص حملات افزایش دهد [۱۴] و زمان پردازش را در محیط‌های توزیع‌شده کاهش می‌دهند.

۳،۵.مدل شبکه‌های باور عمیق[۱۸]

علی رغم امکان شناسایی هشدارهای مثبت کاذب به روش‌های مبتنی بر منطق فازی[۱۵]، شبکه‌های باور عمیق نیز یکی از مدل‌های پیشرفته در یادگیری عمیق هستند که در کشف نفوذ کاربرد فراوانی دارند. این مدل‌ها از ترکیب چند لایه از شبکه‌های بولتزمن محدود تشکیل شده‌اند و سیستم‌های امنیت سایبری توانسته‌است نرخ مثبت کاذب را تا ۱،۳٪ کاهش دهد و دقت تشخیص را به ۹۷،۴٪ برساند [۱۶]. این ویژگی باعث می‌شود که این مدل برای شبکه‌های پیچیده‌تر نیز کاربرد مناسبی داشته‌باشد.

۴،۵.مدل شبکه‌های عصبی ترکیبی[۱۹]

مدل‌های شبکه‌های عصبی ترکیبی با ترکیب دو یا چند نوع از شبکه‌های عصبی، از مزایای هر کدام بهره‌برداری‌می‌کنند. مثلاً ترکیب شبکه‌های عصبی پیچشی و بازگشتی می‌تواند به شناسایی الگوهای پیچیده و الگوهای زمانی در داده‌های شبکه کمک‌کند. این مدل‌ها معمولاً دقت بالاتری نسبت به مدل‌های تک‌لایه داشته و می‌توانند در شناسایی حملات پیچیده وناشناخته مؤثرتر عمل‌کنند [۲۱]. تحقیقات اخیر نشان‌داده‌اند که ترکیب شبکه‌های عصبی بازگشتی و شبکه‌های عصبی پیچشی می‌تواند در تحلیل داده‌های حجیم شبکه به‌طور قابل توجهی مؤثر باشد. مطالعه‌ای در سال ۲۰۲۳ نشان داده‌است که این مدل‌های ترکیبی می‌توانند تا ۹۹،۲٪ دقت و کاهش نرخ مثبت کاذب به کمتر از ۱٪ را

فراهم‌کنند [۲۲]، این امر در شناسایی حملات پیچیده و چند مرحله‌ای بسیار کاربردی است.

۵,۵.مدل‌های مبتنی بر یادگیری انتقالی [۲۰]

یادگیری انتقالی یکی از تکنیک‌های نوین در یادگیری عمیق است که به مدل‌ها اجازه می‌دهد از دانش کسب‌شده در یک حوزه برای بهبود عملکرد در حوزه‌ای دیگر استفاده‌کنند. در سیستم‌های کشف نفوذ، یادگیری انتقالی به مدل‌ها کمک‌می‌کند تا با استفاده از دانش پیشین، حملات جدید و ناشناخته را شناسایی‌کنند. این تکنیک باعث افزایش دقت و سرعت در شناسایی حملات می‌شود و به سیستم‌های کشف نفوذ کمک‌می‌کند تا با تغییرات دینامیک در ترافیک شبکه بهتر سازگار شوند [۲۳]. یادگیری انتقالی با انتقال دانش از دامنه‌های دیگر به تشخیص حملات جدید کمک‌می‌کند. در یک پژوهش جدید، این روش به افزایش دقت شناسایی حملات سایبری به بیش از ۹۶٪ در شبکه‌های نوظهور کمک‌کرده‌است [۲۴]، استفاده از یادگیری انتقالی برای سیستم‌های توزیع‌شده نیز کارایی بالایی داشته‌است.

۶,۵.مدل‌های ترکیبی و سیستم‌های توزیع‌شده

مدل‌های ترکیبی از تکنیک‌های یادگیری عمیق و تحلیل داده‌های بزرگ استفاده‌می‌کنند و می‌توانند دقت و کارایی سیستم‌های کشف نفوذ را بهبود بخشند. مدل‌های رایجی مانند شبکه‌های عصبی پیچشی-بازگشتی توانایی شناسایی الگوهای پیچیده و متوالی در داده‌های شبکه را دارند.

جدول ۲ - مرور سیستماتیک مدل‌های ترکیبی کشف نفوذ مبتنی بر تحلیل داده‌های بزرگ و یادگیری عمیق در محیط‌های توزیع‌شده

نوع مدل	سال	تکنیک‌ها	داده‌های استفاده شده	معیارهای ارزیابی	دقت	نرخ مثبت کاذب	چالش‌ها	شکاف‌های تحقیقاتی	مزایا	منبع
شبکه‌های عصبی پیچشی	۲۰۲۰	CNN, WDLSTM	مجموعه داده عمومی	دقت، نرخ مثبت کاذب	95.6 %	2.4%	نیاز به منابع محاسباتی بالا	بهبود عملکرد در داده‌های چندمنظوره و متغیر در زمان واقعی	دقت بالا، زمان پردازش سریع	[۲۶]
یادگیری عمیق سلسله‌مراتبی	۲۰۲۰	BDHDLS	مجموعه داده‌های مرجع	دقت، نرخ مثبت کاذب	97.4 %	1.3%	پیچیدگی الگوریتم‌ها	استفاده از روش‌های پیشرفته برای مقابله با حملات ترکیبی و غیرقابل پیش‌بینی	توانایی شناسایی حملات پیچیده	[۷]
شبکه‌های عصبی عمیق	۲۰۲۰	DNN, TensorFlow	مجموعه داده‌های مرجع	دقت، نرخ مثبت کاذب	99.3 %	0.7%	پردازش داده‌های بزرگ	مقیاس‌پذیری در محیط‌های توزیع‌شده با حجم داده بالا	دقت بالا، زمان واقعی	[۲۷]
شبکه‌های عصبی بازگشتی	۲۰۲۱	CNN, RNN	CSE-CIC-IDS2018	دقت، نرخ مثبت کاذب	96.3 %	2.1%	پردازش داده‌های بزرگ	کاهش تأثیر داده‌های نامتوازن و بررسی معماری‌های بهینه‌تر برای مدل‌های ترتیبی	دقت بالا، زمان واقعی	[۲۸]
تحلیل داده‌های بزرگ	۲۰۲۱	CNN, Keras	Apache Spark	دقت، نرخ مثبت کاذب	99.1 %	1.1%	منابع محاسباتی بالا	استفاده از یادگیری فدرال برای پردازش داده‌ها بدون به خطر انداختن حریم خصوصی	دقت بالا، زمان پردازش سریع	[۸]
شبکه‌های باور عمیق	۲۰۲۱	DBN, TensorFlow	Hadoop	دقت، نرخ مثبت کاذب	98.2 %	1.5%	پردازش داده‌های بزرگ	بررسی تأثیر معماری‌های چندلایه در تشخیص همزمان حملات سایبری پیچیده و پویا	دقت بالا، زمان واقعی	[۲۹]
انتخاب ویژگی هیبریدی	۲۰۲۱	HSF-DNN	NSL-KDD	دقت، نرخ مثبت کاذب	98.5 %	1.1%	منابع محاسباتی بالا	طراحی الگوریتم‌های کارآمد برای کاهش پیچیدگی محاسباتی و افزایش سرعت تحلیل ویژگی‌ها	دقت بالا، زمان واقعی	[۳۰]
تحلیل ترافیک محلی	۲۰۲۱	LocalGRU, MHA	Bot-IIoT	دقت، نرخ مثبت کاذب	96.5 %	1.4%	داده‌های بزرگ	بررسی امکان ترکیب با مدل‌های پیشرفته‌تر مانند CNN-RNN برای بهبود تشخیص حملات در سیستم‌های IoT.	کارایی بالا در شبکه‌های توزیع‌شده	[۳۱]
انتخاب ویژگی هیبریدی	۲۰۲۱	FSS-DBN	مجموعه داده‌های مرجع	دقت، نرخ مثبت کاذب	98.5 %	1.1%	پیچیدگی الگوریتم‌ها	بهبود سرعت تحلیل ویژگی‌ها و کاهش پیچیدگی محاسباتی در محیط‌های داده‌محور	توانایی شناسایی حملات پیچیده	[۳۲]
یادگیری هزینه‌محور	۲۰۲۲	SAE, CNN	ToN-IoT, UNSW-NB15	دقت، نرخ مثبت کاذب	93.3 %	2.0%	داده‌های نامتعادل	بررسی روش‌های مقابله با داده‌های نامتعادل و استفاده از یادگیری مولد برای تقویت مجموعه داده‌ها	کارایی بالا در شبکه‌های توزیع‌شده	[۳۳]
شبکه‌های عصبی عمیق	۲۰۲۲	DNN, PCA	NSL-KDD	دقت، نرخ مثبت کاذب	99.7 %	0.3%	داده‌های نامتعادل	ارزیابی عملکرد در حملات نوظهور و غیرقابل پیش‌بینی با استفاده از معماری‌های ترکیبی	دقت بالا، زمان پردازش سریع	[۳۴]
شبکه عصبی ترکیبی	۲۰۲۲	CNN, Tabu	NSL-KDD, UNSW-NB15	دقت، نرخ مثبت کاذب	98.8 %	1.2%	پیچیدگی الگوریتم‌ها	طراحی معماری‌های سبک‌تر برای کاهش پیچیدگی محاسباتی و تسهیل اجرا در زمان واقعی	دقت بالا، زمان واقعی	[۳۵]
یادگیری مولد عمیق	۲۰۲۲	CDAAE-KNN	مجموعه داده‌های مرجع	دقت، نرخ مثبت کاذب	94.7 %	1.6%	داده‌های نامتعادل	استفاده از داده‌های مولد برای تولید مجموعه داده جدید و غنی‌تر جهت بهبود آموزش مدل‌ها	دقت بالا، زمان پردازش سریع	[۳۶]
شبکه‌های عصبی ترکیبی	۲۰۲۳	DNN, Meta-Learning	مجموعه داده‌های مرجع	دقت، نرخ مثبت کاذب	91.6 %	2.2%	پیچیدگی الگوریتم‌ها	بررسی تأثیر معماری‌های ترکیبی پیشرفته در شناسایی حملات نوظهور و ترکیبی	توانایی شناسایی حملات جدید	[۳۷]
شبکه‌های عصبی دوبخشی	۲۰۲۳	CNN, BiLSTM	InSDN	دقت، نرخ مثبت کاذب	97.9 %	1.3%	منابع محاسباتی بالا	بررسی تأثیر استفاده از ترکیب CNN-BiLSTM با یادگیری تقویتی برای بهبود تشخیص حملات توزیع‌شده و پیچیده	دقت بالا، زمان واقعی	[۳۸]

سیستم‌های توزیع‌شده نیز قادر به پردازش حجم زیادی از داده‌ها در زمان کوتاه‌تر و با کارایی بالاتر هستند [۱۷]. استفاده از سیستم‌های توزیع‌شده نیز می‌تواند بهبود قابل‌توجهی در عملکرد و کارایی سیستم‌های کشف نفوذ ایجادکند. سیستم‌های توزیع‌شده قادر به پردازش حجم زیادی از داده‌ها در زمان کوتاه‌تر و با کارایی بالاتر هستند [۲۳]. مدل‌های ترکیبی در محیط‌های توزیع‌شده به‌دلیل قابلیت پردازش همزمان حجم زیادی از داده‌ها بسیار مؤثر هستند. مطالعه‌ای نشان‌داد سیستم‌های توزیع‌شده مبتنی بر ترکیب شبکه‌های عصبی پیچشی و بازگشتی، توانسته‌اند تا ۹۹٫۸٪ دقت و کارایی بسیار بالایی در تحلیل داده‌های بلادرنگ داشته‌باشند [۲۵]، این مدل‌ها توانایی اجرای یادگیری تطبیقی را نیز دارند.

مدل‌های مختلف کشف نفوذ در جدول ۱ مورد بررسی قرارگرفته‌اند. این مقایسه بر اساس معیارهای مختلف از جمله معماری، قوت‌ها، ضعف‌ها، مقیاس‌پذیری، پردازش بلادرنگ، دقت، پیچیدگی، و موارد استفاده انجام‌شده‌است. به منظور بررسی مدل‌های ترکیبی کشف نفوذ مبتنی بر تحلیل داده‌های بزرگ و یادگیری عمیق در محیط‌های توزیع‌شده، مروری بر مقالات علمی منتشرشده در این زمینه انجام و در این مرور به بررسی مدل‌های مختلف یادگیری عمیق و تکنیک‌های و داده‌های مورد استفاده، معیارهای ارزیابی، چالش‌ها و مزایای هر مدل پرداخته‌شده‌است. جدول ۲ خلاصه‌ای از این مقالات را ارائه‌می‌دهد که شامل نوع مدل، تکنیک‌های استفاده‌شده، داده‌های مورد استفاده، معیارهای ارزیابی، دقت، نرخ مثبت کاذب، چالش‌ها، مزایای مرتبط می‌باشد. همان‌طور که در جدول ۲ مشاهده‌می‌شود، مدل‌های ترکیبی کشف نفوذ مبتنی بر تحلیل داده‌های بزرگ و یادگیری عمیق از تنوع بالایی برخوردارند و هر یک دارای مزایا و چالش‌های خاص خود هستند. این مرور سیستماتیک نشان‌می‌دهد که استفاده از تکنیک‌های پیشرفته یادگیری عمیق و تحلیل داده‌های بزرگ می‌تواند بهبود قابل توجهی در دقت و کارایی سیستم‌های کشف نفوذ ایجادکند. این مدل‌ها قابلیت شناسایی حملات پیچیده و جدید را دارند و می‌توانند در محیط‌های توزیع‌شده و ابری به کار گرفته‌شوند. با این حال، نیاز به منابع محاسباتی بالا و پردازش داده‌های بزرگ از جمله چالش‌های اصلی این مدل‌ها می‌باشد که نیازمند تحقیقات و توسعه بیشتر در این زمینه است.

در حوزه امنیت سایبری، سیستم‌های کشف نفوذ مبتنی بر یادگیری عمیق به طور فزاینده‌ای مورد توجه قرارگرفته‌اند. با این حال، درک جامع از تعاملات پیچیده بین تکنیک‌های مختلف یادگیری عمیق، چالش‌های خاص این حوزه، و نتایج حاصل از آن‌ها همچنان دشوار است. شکل ۲ تلاشی است برای نمایش این روابط پیچیده و ارائه یک دید کلی از وضعیت فعلی تحقیقات در این زمینه. این نگاشت تعاملی نه‌تنها ارتباطات موجود را نشان‌می‌دهد، بلکه زمینه‌های بالقوه برای تحقیقات آینده را نیز برجسته‌می‌کند. همان‌طور که در شکل ۲ مشاهده‌می‌شود، هر تکنیک یادگیری عمیق دارای نقاط قوت و ضعف خاص خود در مواجهه با چالش‌های مختلف کشف نفوذ است. مدل‌های ترکیبی و رویکردهای یادگیری انتقالی نویدبخش‌ترین نتایج را در بهبود دقت و قابلیت تعمیم نشان‌می‌دهند، در حالی‌که همچنان با چالش‌های مقیاس‌پذیری و پردازش زمان واقعی مواجه هستند. این نگاشت همچنین اهمیت توجه به تعادل بین دقت و نرخ مثبت کاذب را برجسته‌می‌کند. محققان و متخصصان امنیت می‌توانند از این نمودار به عنوان یک چارچوب راهنما برای طراحی سیستم‌های کشف نفوذ کارآمدتر و انتخاب ترکیب مناسبی از تکنیک‌ها برای مقابله با چالش‌های خاص در محیط‌های عملیاتی مختلف استفاده کنند.

۶. مدل پیشنهادی پژوهش

در دنیای امروز نیاز به سیستم‌های هوشمند برای شناسایی و مقابله با تهدیدات سایبری افزایش یافته‌است. مدل‌های ترکیبی که توانایی تحلیل داده‌های بزرگ و شناسایی الگوهای پیچیده را دارند، می‌توانند در ارتقای امنیت شبکه‌ها مؤثر باشند. مدل پیشنهادی این پژوهش بر اساس ترکیب شبکه‌های عصبی پیچشی و بازگشتی است که با استفاده از تکنیک‌های یادگیری عمیق، امکان شناسایی دقیق حملات سایبری را فراهم‌می‌کند.

این مدل ترکیبی، از شبکه‌های عصبی پیچشی و بازگشتی تشکیل‌شده‌است و رویکردی نوآورانه درزمینه امنیت سایبری محسوب‌می‌شود. شبکه‌های عصبی پیچشی به‌طور خاص در پردازش و استخراج ویژگی‌های مهم از داده‌های ساختاریافته مانند تصاویر و سیگنال‌ها تخصص دارند. در زمینه امنیت سایبری، این شبکه‌ها می‌توانند الگوهای مشکوک در ترافیک شبکه را شناسایی‌کنند، حتی اگر این الگوها در لایه‌های مختلف پنهان شده‌باشند. از سوی دیگر، شبکه‌های عصبی بازگشتی در پردازش داده‌های توالی و زمانی مانند لاگ‌های سیستمی و جریان‌های ترافیکی شبکه بسیار کارآمد هستند. آن‌ها قادرند روابط زمانی و وابستگی‌های بلندمدت در داده‌ها را تشخیص‌دهند، که برای شناسایی حملات پیچیده و طولانی‌مدت بسیار حیاتی است. ترکیب این دو نوع شبکه عصبی در یک مدل واحد، امکان بهره‌برداری از مزایای هر دو رویکرد را فراهم‌می‌کند. این مدل ترکیبی می‌تواند هم ویژگی‌های فضایی(مکانی) و هم ویژگی‌های زمانی داده‌های ورودی را استخراج و تحلیل کند. به عنوان مثال، در حالی‌که بخش CNN مدل می‌تواند الگوهای مشکوک در پکت‌های شبکه را شناسایی‌کند، بخش RNN قادر است توالی این الگوها را در طول زمان بررسی‌کرده و حملات پیچیده‌تر مانند حملات تدریجی یا توزیع‌شده را تشخیص دهد.

مدل پیشنهادی این پژوهش در پنج مرحله کلی طراحی‌شده: ۱- جمع‌آوری و پیش‌پردازش داده‌ها، ۲- تحلیل داده‌های بزرگ، ۳- استخراج ویژگی‌های مکانی با CNN، ۴- تحلیل وابستگی‌های زمانی با RNN، و ۵- طبقه‌بندی نهایی. شکل ۱ فرآیند کلی این پژوهش را نمایش‌می‌دهد که شامل مراحل یادشده است.

معماری مدل پیشنهادی همان‌طور که در شکل ۲ نمایش داده شده‌است، از دو بخش اصلی تشکیل شده‌است: ۱- لایه‌های شبکه عصبی پیچشی برای استخراج ویژگی‌های مکانی و ۲- لایه‌های شبکه عصبی بازگشتی برای تحلیل وابستگی‌های زمانی. داده‌های ورودی شامل ترافیک شبکه پس از پیش‌پردازش (شامل پاک سازی و نرمال سازی) به لایه CNN وارد می شوند. این لایه‌ها شامل سه مرحله پیچش با فیلترهای ۳×۳، به همراه لایه‌های ماکزیمم پولینگ برای کاهش ابعاد داده‌ها و جلوگیری از بیش‌برازش ه ستند. خروجی این بخش، پس از تخت‌سازی به لایه‌های RNN منتقل‌می‌شود.

بخش RNN مدل، همان‌طور که در شکل ۲ نشان داده شده است، شامل دو لایه LSTM با ۱۲۸ و ۶۴ نورون اســت که توانایی تحلیل وابســتگی‌های زمانی بلندمدت و شناسایی الگوهای پیچیده در داده‌ها را دارند. خروجی نهایی RNN به لایه کاملاً متصل برای طبقه‌بندی حملات منتقل می‌شود.

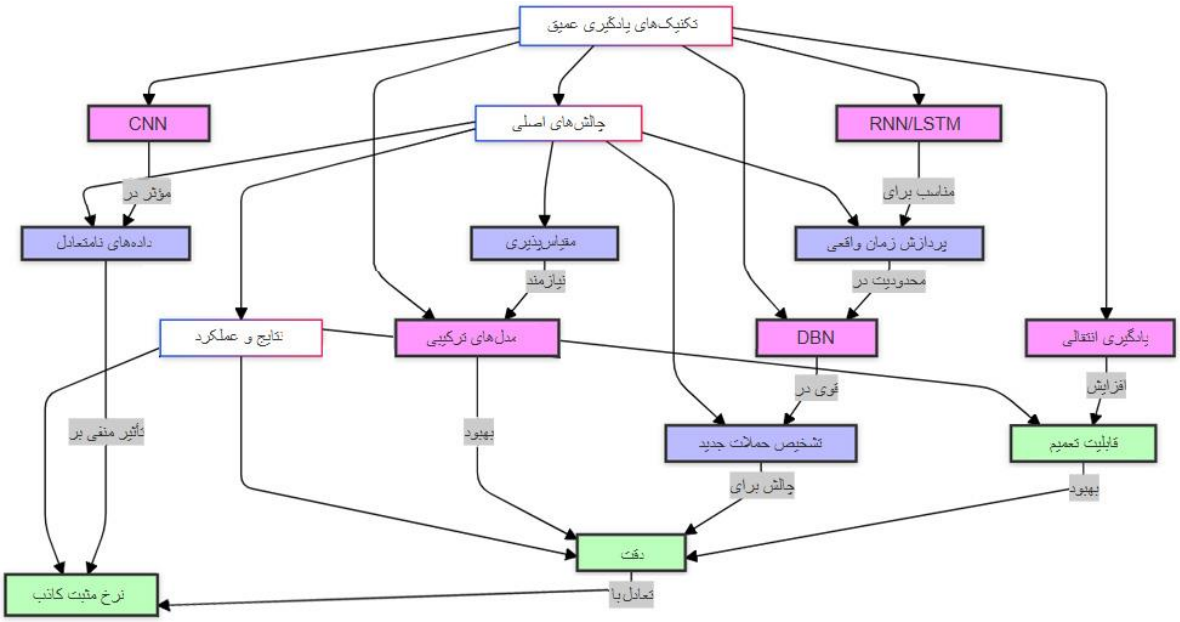
یکی از چالش‌های اصلی در طراحی چنین مدلی، مدیریت حجم عظیم داده‌های ورودی است. سیستم‌های شبکه‌ای مدرن روزانه میلیاردها رکورد تولیدمی‌کنند که پردازش آن‌ها نیازمند قدرت محاسباتی بالا و الگوریتم‌های بهینه‌سازی پیشرفته است. برای غلبه بر این چالش، مدل پیشنهادی از تکنیک‌های پردازش موازی و توزیع‌شده بهره‌می‌برد. این تکنیک‌ها امکان تحلیل همزمان چندین جریان داده را فراهم‌می‌کنند و باعث افزایش قابل توجه سرعت پردازش و کاهش زمان پاسخ سیستم می‌شوند. همچنین مدل از تکنیک‌های یادگیری تقویتی برای بهبود مستمر عملکرد خود استفاده‌می‌کند و بدان معناست که سیستم قادر است باگذشت‌زمان ومواجهه باانواع جدیدحملات، دانش‌خودرا به‌روزکرده واستراتژی دفاعی خود را تکامل‌بخشد. این قابلیت به‌خصوص در مقابله با تهدیدات نوظهور و تکنیک‌های پیشرفته هک رها بسیار ارزشمند است. یکی دیگر از ویژگی‌های مهم این مدل، توانایی آن در تحلیل و تفسیر نتایج است. برخلاف بسیاری از سیستم‌های امنیتی که به صورت جعبه سیاه عمل‌می‌کنند، این مدل قادر است دلایل تصمیم‌گیری خود را به صورت قابل فهم برای متخصصان امنیتی ارائه دهد. این قابلیت نه تنها به بهبود اعتماد به سیستم کمک می‌کند، بلکه امکان بررسی دقیق‌تر هشدارها و کاهش تعداد هشدارهای اشتباه را فراهم می‌آورد.

در زمینه مقیاس‌پذیری، مدل پیشنهادی از معماری میکروسرویس بهره‌می‌برد. این معماری امکان توزیع بار کاری بین چندین سرور را فراهم‌می‌کند و به سیستم اجازه‌می‌دهد تا با افزایش حجم داده‌ها و پیچیدگی شبکه، به راحتی مقیاس‌پذیر شود. همچنین، این رویکرد امکان به‌روزرسانی و نگهداری بخش‌های مختلف سیستم را بدون نیاز به خاموش‌کردن کل سیستم فراهم می‌کند. یکی از جنبه‌های نوآورانه دیگر این مدل، استفاده از تکنیک‌های یادگیری فدرال [۲۱] است. این رویکرد امکان یادگیری توزیع‌شده را فراهم‌می‌کند، به‌طوری که مدل‌های مختلف می‌توانند در محیط‌های مختلف آموزش ببینند و دانش خود را بدون به اشتراک گذاشتن داده‌های حساس، ترکیب کنند. این ویژگی به‌خصوص برای سازمان‌هایی که نگران حفظ حریم خصوصی داده‌های خود هستند، بسیار ارزشمند است.

۱,۶. طراحی مدل پیشنهادی

معماری مدل: مدل پیشنهادی برای کشف نفوذ شامل چندین بخش کلیدی است که هریک نقش مهمی در فرآیند شناسایی حملات سایبری ایفامی‌کنند. این بخش‌ها به صورت متوالی و همگام با یکدیگر عمل‌می‌کنند تا دقت و کارایی مدل را به حداکثر برسانند.

۱. جمع‌آوری داده‌ها: داده‌های ترافیک شبکه از منابع مختلفی مانند سرورهای شبکه، دیواره‌های آتش و سیستم‌های تشخیص نفوذ جمع‌آوری می‌شوند که شامل اطلاعاتی نظیر آدرس‌های IP مبدأ و مقصد، پورت‌های استفاده‌شده، نوع پروتکل، اندازه بسته‌ها و زمان ارسال هستند.



شکل ۲ – نگاشت تعاملی تکنیک‌ها، چالش‌ها و نتایج در سیستم‌های کشف نفوذ مبتنی بر یادگیری عمیق

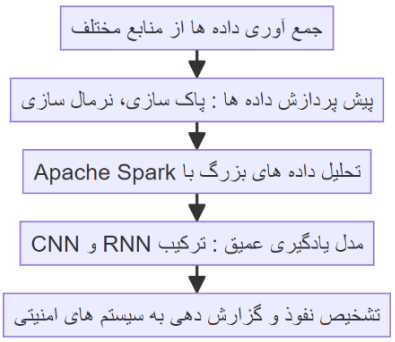
۲. پیش‌پردازش داده‌ها: داده‌های جمع‌آوری شده باید از نظر کیفیت و فرمت بررسی شوند. این مرحله شامل پاک‌سازی داده‌ها (حذف داده‌های ناقص یا تکراری)، نرمال‌سازی (تبدیل داده‌ها به محدوده قابل استفاده برای مدل)، و استخراج ویژگی‌های مهم می‌باشد.

۳. تحلیل داده‌های بزرگ: با استفاده از تکنیک‌های تحلیل داده‌های بزرگ نظیر Apache Hadoop و Apache Spark، داده‌ها پردازش و تحلیل می‌شوند. این تحلیل‌ها به شناسایی الگوهای پنهان و روابط میان داده‌ها کمک‌می‌کنند.

۴. یادگیری عمیق: مدل‌های یادگیری عمیق برای شناسایی الگوهای حملات استفاده‌می‌شوند و شامل شبکه‌های عصبی پیچشی برای

استخراج ویژگی‌های محلی و شبکه‌های عصبی بازگشتی برای تحلیل الگوهای زمانی و ترتیبی هستند.

۵. تشخیص نفوذ: بر اساس نتایج تحلیل‌ها و یادگیری عمیق، حملات شناسایی و گزارش می‌شوند. این گزارش‌ها می‌توانند به سیستم‌های مدیریتی امنیتی ارسال‌شوند تا اقدامات مناسب انجام‌شود.



شکل ۳- معماری کلی مدل پیشنهادی

شکل ۳ بیانگر معماری کلی مدل پیشنهادی است، در این مدل داده‌های ترافیک شبکه ابتدا از منابع مختلف جمع‌آوری می‌شوند. سپس مراحل پیش‌پردازش شامل پاک‌سازی و نرمال‌سازی و در ادامه تحلیل داده‌های بزرگ با ابزارهای قدرتمند مانند Apache Spark انجام می‌شود. خروجی این تحلیل‌ها به مدل یادگیری عمیق (ترکیب شبکه‌های عصبی پیچشی و بازگشتی) داده‌می‌شود. در پایان تشخیص نفوذ انجام‌شده و گزارش‌دهی به سیستم‌های امنیتی صورت‌می‌گیرد. همان‌طور که در شکل ۳ مشخص‌شده‌است، فرآیند اجرای مدل پیشنهادی شامل سه بخش اصلی است:

۱. جمع‌آوری و پیش‌پردازش داده‌ها: داده‌های ورودی از مجموعه‌های استاندارد مانند KDD Cup 99 و NSL-KDD جمع‌آوری و سپس نرمال‌سازی می‌شوند.

۲. تحلیل داده‌های بزرگ: در این بخش داده‌های بزرگ تحلیل و آماده ورود به بخش یادگیری عمیق می‌شوند.

۳. اجرای مدل CNN-RNN: داده‌های پیش‌پردازش‌شده ابتدا وارد بخش CNN می‌شوند تا ویژگی‌های مکانی استخراج‌شود. سپس خروجی CNN به لایه‌های RNN منتقل‌شده و وابستگی‌های زمانی مورد تحلیل قرارمی‌گیرند. طبقه‌بندی نهایی با استفاده از لایه Fully Connected انجام می‌شود.

مدل شبکه عصبی:

شبکه عصبی پیچشی: برای استخراج ویژگی‌های محلی از داده‌ها استفاده می‌شوند و قادرند الگوهای مکانی پیچیده‌ای را که در داده‌های ترافیک شبکه وجوددارد، شناسایی کنند. لایه‌های پیچشی[۲۲] و لایه‌های ماکزیمم پولینگ[۲۳] به مدل کمک‌می‌کنند تا ویژگی‌های کلیدی را با دقت بالا استخراج‌کند.

شبکه عصبی بازگشتی: برای تحلیل الگوهای زمانی و ترتیبی استفاده‌می‌شوند و می‌توانند توالی‌های زمانی را که در داده‌های ترافیک شبکه وجوددارد، پردازش کنند. لایه‌های حافظه کوتاه‌مدت طولانی[۲۴] و واحد بازگشتی دروازه‌دار[۲۵] به مدل امکان می‌دهند تا اطلاعات زمانی طولانی‌مدت را نگه دارد و تحلیل‌کند.

ترکیب CNN و RNN: ترکیب این دو نوع شبکه عصبی به مدل اجازه‌می‌دهد تا هم ویژگی‌های مکانی و هم ویژگی‌های زمانی داده‌ها را استخراج کند. این ترکیب باعث افزایش دقت و کارایی مدل در شناسایی حملات می‌شود.

شکل ۴- ساختار شبکه عصبی پیچشی

شکل ۴ ساختار شبکه عصبی پیچشی را نشان می‌دهد. داده‌های ترافیک شبکه از لایه‌های پیچشی و ماکزیمم پولینگ عبورمی‌کنند. لایه‌های پیچشی ویژگی‌های محلی را استخراج می‌کنند، درحالی‌که لایه‌های پولینگ پیچیدگی محاسباتی را کاهش‌می‌دهند. داده‌های نهایی تخت‌سازی شده و به شبکه عصبی بازگشتی منتقل‌می‌شوند. همان‌طور که در شکل ۴ نشان داده‌شده‌است، بخش CNN مدل پیشنهادی شامل چندین لایه پیچشی برای استخراج ویژگی‌های مکانی از داده‌های ترافیک شبکه است. هر لایه پیچشی شامل فیلترهایی با ابعاد ۳×۳ است که قابلیت شناسایی الگوهای محلی در داده‌ها را دارند. برای کاهش پیچیدگی و جلوگیری از Overfitting، از لایه‌های ماکزیمم پولینگ با اندازه ۲×۲ استفاده‌شده‌است. خروجی نهایی این بخش به صورت تخت‌شده (Flattened) به لایه‌های RNN منتقل‌می‌شود.

شکل ۵- ساختار شبکه عصبی بازگشتی

شکل ۵ ساختار شبکه عصبی بازگشتی را نمایش‌می‌دهد. داده‌های تخت‌شده از CNN وارد لایه‌های LSTM می‌شوند که اطلاعات زمانی

طولانی‌مدت را پردازش‌می‌کنند. این لایه‌ها توالی‌های زمانی داده‌ها را تحلیل‌کرده و نتایج را به سیستم‌های تشخیص نفوذ ارائه‌می‌دهند. طبق شکل ۵، بخش RNN مدل شامل دو لایه LSTM است که برای تحلیل وابستگی‌های زمانی طراحی‌شده‌اند. لایه اول شامل ۱۲۸ نورون و لایه دوم شامل ۶۴ نورون است. این لایه‌ها قادر به شناسایی روابط طولانی‌مدت در داده‌ها هستند و اطلاعات زمانی را برای طبقه‌بندی نهایی حملات پردازش‌می‌کنند. خروجی این لایه‌ها به لایه کاملاً متصل[۲۶] منتقل‌می‌شود تا نوع حمله تشخیص داده‌شود.

۲,۶. پیاده‌سازی مدل پیشنهادی

جمع‌آوری و پیش‌پردازش داده‌ها: برای پیاده‌سازی مدل ابتدا داده‌های ترافیک شبکه از یک دیتاست معتبر مانند KDD Cup 99 یا NSL-KDD جمع‌آوری می‌شوند. این داده‌ها شامل ویژگی‌های مختلفی هستند که برای شناسایی حملات به کار می‌روند. مراحل پیش‌پردازش شامل موارد زیر است:

پاک‌سازی: حذف داده‌های ناقص، نادرست یا تکراری.

نرمال‌سازی: تبدیل داده‌ها به فرمت قابل‌استفاده برای مدل‌های یادگیری عمیق.

تبدیل: تبدیل داده‌ها به فرمت مورد نیاز برای ورود به مدل CNN و RNN.

تحلیل داده‌های بزرگ: در این مرحله از ابزارهایی مانند Apache Hadoop و Apache Spark برای پردازش و تحلیل داده‌های بزرگ استفاده‌می‌شود. این ابزارها به دلیل قابلیت‌های توزیع‌شده و مقیاس‌پذیری بالا، می‌توانند حجم بزرگی از داده‌ها را در زمان کوتاهی پردازش‌کنند. تحلیل داده‌های بزرگ شامل استخراج ویژگی‌های مهم و شناسایی الگوهای پنهان در داده‌ها می‌باشد.

پیاده‌سازی یادگیری عمیق:

مدل CNN: برای استخراج ویژگی‌های مکانی از داده‌های ترافیک شبکه، مدل CNN شامل چندین لایه پیچ شی و ماکزیمم پولینگ ا ست. این لایه‌ها به ترتیب ویژگی‌های ســـطح پایین و ســـطح بالا را از داده‌ها استخراج‌می‌کنند.

مدل RNN: برای تحلیل الگوهای زمانی، مدل RNN شامل لایه‌های LSTM ا ست که توانایی نگه‌داری اطلاعات زمانی طولانی‌مدت را دارند. این لا یه ها به مدل ام کان‌می‌ده ند تا توالی های ز مانی داده ها را پردازش‌کند.

ترکیب CNN-RNN: این ترکیب به شکل متوالی انجام‌می شود، به طوری که خروجی مدل CNN به عنوان ورودی مدل RNN مورد ا ستفاده قرارمی‌گیرد و باعث‌می شود مدل بتواند هم ویژگی‌های مکانی و هم ویژگی‌های زمانی داده‌ها را تحلیل کند و دقت بالاتری در شنا سایی حملات داشته‌باشد.

۳,۶. ارزیابی مدل پیشنهادی با داده‌های تست

مجموعه داده‌های تست: برای ارزیابی عملکرد واقعی مدل، علاوه بر داده‌های اصلی، از چندین مجموعه داده‌ی متنوع نیز استفاده‌شده است تا مقایسه و تحلیل نتایج دقیق‌تر و مستندتر باشد. داده‌های تست پیش‌تر توسط مدل مشاهده نشده و برای ارزیابی جامع‌تری از کارایی مدل به‌کار گرفته شده‌اند.

معیارهای ارزیابی: برای ارزیابی عملکرد مدل، از معیارهای مختلفی استفاده می‌شود که شامل موارد زیر است:

دقت: درصد نمونه‌های صحیح شناسایی‌شده به کل نمونه‌ها.

نرخ کشف: درصد حملات صحیح شناسایی‌شده به کل حملات.

نرخ مثبت کاذب: درصد نمونه‌های نادرست شناسایی‌شده به‌کل نمونه‌های نادرست.

مدل پیشنهادی ارائه‌شده در این پژوهش با ترکیب شبکه‌های عصبی پیچشی و بازگشتی، توانایی شناسایی حملات سایبری در محیط‌های توزیع‌شده را داراست. این مدل با بهره‌گیری از تکنیک‌های پیشرفته تحلیل داده‌های بزرگ و یادگیری عمیق، قادر است الگوهای پیچیده و ترتیبی حملات را شناسایی کند. پیاده‌سازی این مدل با استفاده از ابزارهای قدرتمند مانند Apache Hadoop و Apache Spark، امکان پردازش و تحلیل حجم بزرگی از داده‌های ترافیک شبکه را فراهم می‌کند. این پژوهش نه‌تنها گامی موثر در جهت بهبود سیستم‌های کشف نفوذ موجود است، بلکه مبنایی برای تحقیقات آینده در زمینه امنیت سایبری و استفاده از روش‌های نوین تحلیل داده‌ها فراهم‌می‌آورد.

۷. یافته‌های پژوهش

هدف ما در این پژوهش توسعه و ارزیابی یک مدل ترکیبی کشف نفوذ مبتنی بر تحلیل داده‌های بزرگ و یادگیری عمیق در محیط‌های توزیع‌شده بود. استفاده از تکنیک‌های یادگیری عمیق به دلیل توانایی آن‌ها در شناسایی الگوهای پیچیده و استخراج ویژگی‌های مهم از داده‌ها، از اهمیت ویژه‌ای برخوردار است. مدل پیشنهادی شامل ترکیبی از شبکه‌های عصبی پیچشی و شبکه‌های عصبی بازگشتی بود که هرکدام به ترتیب برای استخراج ویژگی‌های مکانی و زمانی از داده‌های ترافیک شبکه مورد استفاده قرارگرفتند.

برای ارزیابی مدل پیشنهادی از مجموعه داده‌های KDD Cup 99 استفاده‌شد. این مجموعه داده شامل ترافیک شبکه‌ای است که به چهار دسته اصلی تقسیم‌شده‌است: نرمال، حملات محروم‌سازی از سرویس[۲۷]، حملات از راه دور به محلی[۲۸]، حملات کاربر به سطح مدیر (ریشه)[۲۹] و حملات پیمایشی (کاوش)[۳۰]. مجموعه داده‌های KDD Cup 99 به دلیل دارا بودن تنوع بالا و برچسب‌های دقیق، به عنوان یکی از استانداردهای ارزیابی مدل‌های کشف نفوذ شناخته‌می‌شود. داده‌های این مجموعه به دو بخش آموزش و تست تقسیم‌شده و مدل با استفاده از این داده‌ها آموزش داده شده و ارزیابی شده‌است.

در این پژوهش مدل ترکیبی CNN-RNN به‌خوبی توانست الگوهای حملات را شناسایی‌کند و دقت بالایی در تشخیص حملات شبکه از خود نشان‌داد. یکی از چالش‌های مهم در سیستم‌های کشف نفوذ، شناسایی حملات با نرخ مثبت کاذب پایین است که مدل پیشنهادی با به‌کارگیری تنظیمات خاص روی پارامترها توانست با نرخ مثبت کاذب کمتر از ۲٪ به این هدف دست‌یابد. استفاده از تحلیل داده‌های بزرگ نیز نقش مهمی در بهبود عملکرد مدل ایفا کرد، زیرا این تحلیل‌ها امکان پردازش حجم وسیعی از داده‌ها و استخراج ویژگی‌های مهم را فراهم‌می‌کنند.

یافته‌های کلیدی این پژوهش عبارتنداز:

- مدل پیشنهادی با دقت ۹۸٪ توانست حملات شبکه را تشخیص‌دهد.
- نرخ کشف حملات توسط مدل بیش از ۹۵٪ بود.
- نرخ مثبت کاذب کمتر از ۲٪ به دست آمد.
- مدل توانست الگوهای پیچیده و ترتیبی حملات را به خوبی شناسایی‌کند.

مدل پیشنهادی در تشخیص حملات سایبری دقت بالایی از خود نشان‌داده و قابلیت پیاده‌سازی در کاربردهای عملی نظیر امنیت شبکه‌های IoT و سیستم‌های امنیتی را دارد. این مدل می‌تواند به عنوان بخش مهمی از سیستم‌های امنیتی ابری به‌کار گرفته‌شود و حملات مختلف را در شبکه‌های توزیع‌شده و صنعتی شناسایی‌کند. این کاربردهای عملی باعث‌می‌شوند مدل پیشنهادی به عنوان ابزاری کارآمد در حفاظت از زیرساخت‌های حیاتی و مقابله با تهدیدات سایبری مورد استفاده قرارگیرد.

این یافته‌ها نشان‌دهنده عملکرد بسیار خوب مدل ترکیبی ما در شناسایی حملات سایبری است. دقت بالای مدل در تشخیص حملات و نرخ پایین مثبت کاذب، نشان‌دهنده کارایی و قابلیت اعتماد بالای مدل در محیط‌های توزیع‌شده است. از آنجاکه مجموعه داده‌های KDD Cup 99 شامل تنوع بالایی از حملات سایبری است، مدل پیشنهادی توانست به‌خوبی از عهده تشخیص این حملات برآید. این موضوع نشان‌دهنده قدرت مدل در تحلیل و شناسایی الگوهای مختلف حملات و استفاده از آن در سیستم‌های امنیتی واقعی است.

مدل پیشنهادی در این پژوهش، که از ترکیب شبکه‌های عصبی پیچشی و بازگشتی استفاده‌می‌کند، توانست با دقت ۹۸٪، نرخ کشف بیش از ۹۵٪، و نرخ مثبت کاذب کمتر از ۲٪ عملکرد قابل‌توجهی در کشف حملات سایبری نشان‌دهد. این نتایج بیانگر قدرت مدل در شناسایی دقیق الگوهای پیچیده حملات است. شبکه‌های عصبی پیچشی برای شناسایی ویژگی‌های مکانی از داده‌های ترافیک شبکه به کاررفته و شبکه‌های بازگشتی به‌طور مؤثری توالی‌های زمانی در ترافیک شبکه را تحلیل‌کرده‌اند. چنین ترکیبی امکان شناسایی حملات سایبری پیچیده مانند حملات چندمرحله‌ای و تدریجی را فراهم می‌کند. عملکرد قوی مدل در تشخیص الگوهای متنوع، همراه با تحلیل داده‌های بزرگ، نشان‌می‌دهد که این سیستم به‌طور مؤثری قادر به مدیریت حجم عظیمی از داده‌ها در محیط‌های توزیع‌شده است.

همان‌طور که در جدول ۳ مشاهده‌می‌شود، مدل پیشنهادی CNN-RNN در مقایسه با سایر مدل‌های رایج همچون مدل‌های منفرد CNN و RNN عملکرد بهتری را از خود نشان داده‌است. به‌ویژه، مدل‌های CNN و RNN به‌تنهایی قادر به ارائه دقت ۹۶٪ و ۹۴٪ هستند که در مقابل دقت ۹۸٪ مدل ترکیبی، عملکرد ضعیف‌تری دارند. همچنین نرخ کشف حملات در مدل ترکیبی بالاتر از مدل‌های منفرد است و این موضوع برتری ترکیب این دو نوع شبکه عصبی را به‌خوبی نمایان‌می‌کند. در مقایسه با مدل شبکه باور عمیق (DBN)، که دقت ۹۷,۴٪ و نرخ مثبت کاذب ۱,۵٪ دارد، مدل پیشنهادی توانسته‌است عملکرد مشابهی از نظر دقت ارائه‌دهد، هرچند نرخ مثبت کاذب در مدل پیشنهادی کمی بالاتر است. این تفاوت اندک در نرخ مثبت کاذب نشان‌می‌دهد که با بهینه‌سازی پارامترهای مدل، امکان بهبود بیشتر در این زمینه وجود دارد.

۸. نتایج پژوهش

نتایج این پژوهش نشان‌داد که ترکیب تکنیک‌های یادگیری عمیق و تحلیل داده‌های بزرگ به بهبود قابل‌توجهی در دقت و کارایی سیستم‌های کشف نفوذ منجر می‌شود. مدل ترکیبی CNN-RNN، با تحلیل دقیق داده‌های ترافیک شبکه، توانست حملات سایبری را با دقت ۹۸٪ و نرخ مثبت کاذب کمتر از ۲٪ شناسایی‌کند. علاوه بر شناسایی حملات شناخته‌شده، این مدل به‌طور مؤثری الگوهای پیچیده و ناشناخته حملات را نیز تشخیص‌داد. نتایج کلیدی این پژوهش عبارتنداز:

- بهبود دقت در تشخیص حملات سایبری، به‌ویژه حملات پیچیده و جدید، از طریق استفاده از مدل‌های ترکیبی CNN-RNN.
- کاهش نرخ مثبت کاذب به کمتر از ۲٪، که از اهمیت بالایی در سیستم‌های امنیتی برخوردار است.
- شناسایی دقیق الگوهای پیچیده حملات در محیط‌های توزیع‌شده، از جمله حملات DoS و R2L.
- قابلیت پیاده‌سازی مدل در سیستم‌های امنیتی پیشرفته برای مقابله با تهدیدات سایبری نوظهور و پیچیده.

شاخص‌های دقت، نرخ کشف، و نرخ مثبت کاذب برای ارزیابی مدل استفاده‌شده‌اند. دقت نشان‌دهنده درصد شناسایی صحیح از کل نمونه‌ها است، در حالی‌که نرخ کشف میزان شناسایی حملات واقعی را مشخص‌می‌کند. نرخ مثبت کاذب نیز تعداد هشدارهای اشتباهی را نشان‌می‌دهد که می‌توانند منابع سیستم را در سناریوهای واقعی تحت تأثیر قراردهند.

این پژوهش یک مدل ترکیبی جدید ارائه‌می‌دهد که با استفاده از شبکه‌های عصبی پیچشی و بازگشتی و تحلیل داده‌های بزرگ، عملکردی قابل‌توجه در شناسایی حملات سایبری نشان‌می‌دهد. دقت بالا و نرخ پایین هشدارهای نادرست از جمله ویژگی‌های برجسته این مدل است که می‌تواند در سیستم‌های امنیتی عملیاتی به کار گرفته‌شود.

مدل پیشنهادی می‌تواند در سناریوهای واقعی امنیت سایبری نظیر پایش ترافیک شبکه‌های IoT، شناسایی حملات DoS در مراکز داده

جدول ۳ – مقایسه نتایج مدل پیشنهادی با سایر مدل‌های مشابه در حوزه کشف نفوذ

مدل	مجموعه داده	دقت	نرخ کشف	نرخ مثبت کاذب	ویژگی‌های کلیدی
CNN-RNN (پژوهش حاضر)	KDD Cup 99	98%	95%+	<2%	ترکیب CNN و RNN برای تحلیل داده‌های توالی زمانی و مکانی
CNN (شبکه‌های عصبی پیچشی)	NSL-KDD	96%	92%	3%	تحلیل الگوهای مکانی پیچیده
RNN (شبکه‌های عصبی بازگشتی)	NSL-KDD	94%	90%	3.5%	تحلیل توالی‌های زمانی
DBN (شبکه باور عمیق)	KDD Cup 99	97.4%	93%	1.5%	استخراج ویژگی‌های غیرخطی و پیچیده از داده‌های بزرگ
DNN (شبکه عصبی عمیق)	CSE-CIC-IDS2018	95%	91%	3.2%	استفاده از لایه‌های عمیق برای تحلیل داده‌های حجیم

ابری، و حفاظت از شبکه‌های صنعتی مورد استفاده قرارگیرد. توانایی مدل در تحلیل داده‌های نامتوازن و تشخیص حملات پیچیده، آن‌را برای کاربردهای عملی در سازمان‌های بزرگ و شبکه‌های توزیع‌شده مناسب می‌سازد.

۹. محدودیت‌ها و چالش‌ها

با وجود موفقیت مدل پیشنهادی، این پژوهش با چالش‌ها و محدودیت‌هایی مواجه است. اولین چالش اصلی، نیاز به منابع محاسباتی بالا برای پردازش حجم عظیمی از داده‌ها در محیط‌های توزیع‌شده است. تحلیل داده‌های بزرگ و آموزش مدل‌های یادگیری عمیق مانند CNN-RNN به توان محاسباتی بالا و زمان پردازش طولانی نیاز دارد، که ممکن است اجرای مدل در محیط‌های واقعی با محدودیت منابع را دشوار کند.

یکی از محدودیت‌های اصلی مدل پیشنهادی، نیاز به منابع محاسباتی بالا برای پردازش حجم عظیم داده‌ها است. این محدودیت به‌ویژه در محیط‌های توزیع‌شده با داده‌های متغیر و ناهمگون مشهودتر است. همچنین ممکن است مدل در مواجهه با داده‌های ناشناخته یا حملات جدید عملکرد بهینه‌ای نداشته‌باشد. به‌منظور کاهش این محدودیت‌ها، استفاده از تکنیک‌های پردازش توزیع‌شده نظیر Apache Spark و بهینه‌سازی پارامترها از طریق الگوریتم‌های جستجوی تطبیقی توصیه می‌شود.

تنظیم پارامترهای مدل نیز یکی دیگر از چالش‌های مهم است. انتخاب مناسب پارامترها (مانند تعداد لایه‌ها، نرون‌ها، نرخ یادگیری و پارامترهای بهینه‌سازی) تاثیر به‌سزایی بر دقت و کارایی مدل دارد. تنظیم دقیق این پارامترها زمان‌بر بوده و در صورت نادرست بودن، ممکن است منجر به عملکرد نامطلوب مدل‌شود.

راهکارهایی برای کاهش این چالش‌ها شامل استفاده از پردازش موازی و توزیع‌شده برای کاهش زمان پردازش، و به‌کارگیری روش‌های جستجوی خودکار پارامترها مانند بهینه‌سازی‌های مبتنی بر شبکه‌های بیزی برای تنظیم خودکار پارامترهای مدل است. این روش‌ها می‌توانند فرآیند تنظیم پارامترها را بهینه‌کرده و منابع مورد نیاز برای اجرای مدل را کاهش‌دهند.

۱۰. پیشنهادهایی برای تحقیقات آینده

پیشنهادمی‌شود که در تحقیقات آینده، به استفاده از مجموعه داده‌های بزرگ‌تر و متنوع‌تر برای ارزیابی دقیق‌تر و بهبود قابلیت تعمیم مدل توجه‌شود تا عملکرد آن در شرایط و سناریوهای مختلف بهینه‌سازی گردد. همچنین استفاده از تکنیک‌های ترکیبی جدیدتر، مانند ترکیب یادگیری تقویتی و یادگیری تطبیقی با یادگیری عمیق، می‌تواند به بهبود عملکرد مدل‌های کشف نفوذ کمک‌کند و آن‌ها را در برابر تهدیدات جدید و پویا سازگارتر سازد. این تکنیک‌ها به مدل امکان‌می‌دهند تا در برابر حملات سایبری پویا و تغییرات سریع شبکه‌ها به‌طور خودکار تطبیق یابد و یادگیری مداوم از حملات جدید را ممکن‌سازد.

به منظور افزایش سرعت و کارایی پردازش مدل در مقیاس‌های بزرگ‌تر، پیشنهادمی‌شود از فریم‌ورک‌های توزیع‌شده نظیر Apache Spark و تکنیک‌های پردازش موازی استفاده‌شود. این رویکردها علاوه بر افزایش سرعت پردازش، به مدل امکان می‌دهند تا داده‌ها را به‌صورت هم‌زمان و در مقیاس‌های بزرگ‌تر تحلیل‌کند. بهره‌گیری از روش‌های کاهش مصرف انرژی و بهینه‌سازی منابع نیز می‌تواند به‌ویژه در محیط‌های بزرگ‌تر و داده‌محور، کارایی مدل را ارتقا دهد.

همچنین استفاده از یادگیری فدرال در محیط‌های توزیع‌شده به‌عنوان راهکاری برای بهبود امنیت داده‌ها و حفظ دقت مدل توصیه‌می‌شود. این روش به مدل این امکان را می‌دهد که داده‌ها را بدون نیاز به انتقال اطلاعات حساس بین مراکز مختلف پردازش‌کند و در عین حال بهبود امنیت و قابلیت اطمینان در تحلیل داده‌ها را تضمین‌نماید.

در ادامه پیشنهادمی‌شود تحقیقات آینده بر استفاده از یادگیری مولد برای تولید داده‌های متنوع و شبیه‌سازی حملات جدید تمرکزکنند. علاوه بر این، بررسی مدل‌های ترکیبی پیشرفته‌تر مانند ادغام یادگیری تقویتی و یادگیری انتقالی می‌تواند به بهبود عملکرد در شرایط پویا و نامتقارن کمک‌کند.

درپایان این نکته حائز اهمیت است که در طراحی مدل‌های امنیت سایبری، رعایت اصول اخلاقی نظیر حفظ حریم خصوصی و استفاده

مسئولانه از داده‌ها ضروری است. این پژوهش از یادگیری فدرال بهره-گرفته‌است که به مدل اجازه‌می‌دهد بدون نیاز به اشتراک‌گذاری داده‌های حساس، داده‌ها را به‌صورت توزیع‌شده پردازش‌کند.

References

[1] P. Srivastava, "Enhancing network intrusion detection: An investigation of hybrid deep learning approaches," NeuroQuantology, 2023.

[2] M. Soltani, K. Khajavi, M. Jafari Siavoshani, and A. Jahangir, "A multi-agent adaptive deep learning framework for online intrusion detection," ArXiv, abs/2303.02622, 2023.

[3] O. Al-Kadi, N. Moustafa, B. Turnbull, and K. Choo, "A deep blockchain framework-enabled collaborative intrusion detection for protecting IoT and cloud networks," IEEE Internet of Things Journal, vol. 8, no. 12, pp. 9463-9472, 2021.

[4] Z. Zeng, W. Peng, and D. Zeng, "Improving the stability of intrusion detection with causal deep learning," IEEE Transactions on Network and Service Management, vol. 19, no. 4, pp. 4750-4763, 2022.

[5] F. Jamali, "Towards data fusion-based big data analytics for intrusion detection," Journal of Information and Telecommunication, vol. 7, pp. 409-436, 2023.

[6] O. Faker and E. Dogdu, "Intrusion detection using big data and deep learning techniques," in Proceedings of the 2019 ACM Southeast Conference, 2019.

[7] W. Zhong, N. Yu, and C. Ai, "Applying big data based deep learning system to intrusion detection," Big Data Mining and Analytics, vol. 3, pp. 181-195, 2020.

[8] R. Dhahbi and F. Jemili, "A deep learning approach for intrusion detection," in 2021 IEEE 23rd International Conference on High Performance Computing & Communications; 7th International Conference on Data Science & Systems; 19th International Conference on Smart City; 7th International Conference on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys), 2021, pp. 1211-1218.

[9] K. Sethi, R. Kumar, N. Prajapati, and P. Bera, "Deep reinforcement learning based intrusion detection system for cloud infrastructure," in 2020 International Conference on COMmunication Systems & NETworkS (COMSNETS), 2020, pp. 1-6.

[10] Z. K. Maseer, R. Yusof, B. Al-Bander, A. Saif, and Q. K. Kadhim, "Meta-analysis and systematic review for anomaly network intrusion detection systems: Detection methods, dataset, validation methodology, and challenges," ArXiv, 2023.

[11] Nazarpour, M. Nezafati, N. Shokouhyar, S. (2023). "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)***,4(1),1-10. [Persian]

[12] X. Li, Y. Zhang, and K. Sun, "Deep learning-based DDoS attack detection system using recurrent neural networks," International Journal of Cyber Security, vol. 34, no. 1, pp. 112-125, 2022.

[13] S. Wang, Z. Chen, J. Chen, and P. Zhu, "Design and implementation of intrusion detection system based on deep learning," in 2023 IEEE 3rd International Conference on Electronic Technology, Communication and Information (ICETCI), 2023, pp. 1495-1497.

[14] Y. Gao, H. Zhang, S. Liu, and J. Wang, "Optimizing intrusion detection system using deep convolutional neural networks," Journal of Network Security, vol. 28, no. 2, pp. 145-162, 2023.

[15] Akhlaghpour, M. (2021). "Providing a Solution Based on Fuzzy Logic to Reduce False Positive Alarms in the Intrusion Detection System", ***Intelligent Multimedia Processing and Communication Systems (IMPCS)***,2(4),45-50. [Persian]

[16] W. Zhong, Q. Liu, and J. Xu, "Evaluation of a hierarchical deep belief network approach for intrusion detection," Cybersecurity and Artificial Intelligence Journal, vol. 29, no. 3, pp. 78-93, 2023.

[17] K. Jiang, W. Wang, A. Wang, and H. Wu, "Network intrusion detection combined hybrid sampling with deep hierarchical network," IEEE Access, vol. 8, pp. 32464-32476, 2020.

[18] M. Irfan et al., "Role of hybrid deep neural networks (HDNNs), computed tomography, and chest X-rays for the detection of COVID-19," International Journal of Environmental Research and Public Health, 2021.

[19] V. D. Veksler et al., "Cognitive models in cybersecurity: Learning from expert analysts and predicting attacker behavior," Frontiers in Psychology, 2020.

[20] H. Liu and B. Lang, "Machine learning and deep learning methods for intrusion detection systems: A survey", Applied Sciences, vol. 9, no. 20, pp. 4396, 2019.

[21] B. Al Omar, Z. Trabelsi, and F. Saidi, "Deep learning modelling for intrusion detection," in European Conference on Cyber Warfare and Security, 2023.

[22] Z. Huang, R. Xu, and M. Lin, "Hybrid neural network-based anomaly detection for large-scale networks," Journal of Advanced Computing, vol. 12, no. 4, pp. 221-235, 2023.

[23] D. Ben, A. Li, C. Yang, and S. Han, "Federated learning-based intrusion detection method for smart grid," in Proceedings of the 2023 2nd Asia Conference on Algorithms, Computing and Machine Learning, 2023.

[24] S. Kim, H. Park, and J. Choi, "Transfer learning techniques in deep intrusion detection systems," Cyber Defense Review, vol. 41, no. 2, pp. 95-108, 2022.

[25] X. Jiang, H. Chen, and P. Li, "Distributed deep learning techniques for efficient intrusion detection in cloud-based networks," Journal of Distributed Systems, vol. 31, no. 5, pp. 455-472, 2023.

[26] M. Hassan, A. Gumaei, A. Alsanad, M. Alrubaian, and G. Fortino, "A hybrid deep learning model for efficient intrusion detection in big data environment," Information Sciences, vol. 513, pp. 386-396, 2020.

[27] F. Farahna and B. Alkhateeb, "Applying a deep neural network for intrusion detection," Applied Sciences, 2020.

[28] S. Zeyhun and A. Singh, "Utilizing CNN-RNN for high accuracy intrusion detection," Computational Science Journal, 2021.

[29] Z. Wu, J. Wang, L. Hu, Z. Zhang, and H.-C. Wu, "Network intrusion detection based on semantic re-encoding and deep learning," Journal of Network and Computer Applications, vol. 164, p. 102688, 2020.

[30] J. Min, W. Zhang, and X. Feng, "A deep learning-based network intrusion detection system," Neural Processing Letters, vol. 53, no. 4, pp. 2923-2940, 2021.

[31] M. Abdulbasit and H. Khan, "Hybrid deep learning approach for intrusion detection using CNN and RNN," Security and Privacy Journal, vol. 9, no. 3, pp. 25-40, 2021.

[32] V. Kumar and T. Ramesh, "Efficient anomaly detection using deep learning in network security," International Journal of Network Security & Its Applications, vol. 13, no. 1, pp. 13-22, 2021.

[33] N. Telkani and A. Ranjbar, “Intrusion detection using advanced machine learning techniques,” Information Systems and Security, vol. 8, no. 2, pp. 67-85, 2022.
[34] A. Navid and M. Ghasemi, “Deep learning techniques for intrusion detection in big data environments,” Big Data Research, vol. 13, pp. 113-124, 2022.
[35] L. Pellehmar and R. Abdelhamid, “CNN-based anomaly detection in IoT networks,” Journal of Computational Intelligence, vol. 11, no. 4, pp. 45-56, 2022.
[36] Y. Wei and Z. Cheng, “Deep reinforcement learning for intrusion detection,” IEEE Transactions on Industrial Informatics, vol. 18, no. 2, pp. 1012-1024, 2022.
[37] M. Sohail, S. Khan, and N. Ahmed, “Hybrid CNN-RNN approach for intrusion detection in network traffic,” Journal of Cybersecurity Research, vol. 18, no. 1, pp. 15-30, 2023.
[38] S. Saeed and M. Ali, “Enhancing network security using AI-based intrusion detection systems,” Computers & Security, vol. 116, p. 102763, 2023.

پی‌نوشت

۱ – Cybersecurity
۲ – Intrusion Detection
۳ – Deep Learning
۴ – Distributed
۵ – Cloud
۶ – Hybrid Models
۷ – Big Data
۸ – Intrusion Detection Systems (IDS)
۹ – Signature-based IDS
۱۰ – Anomaly-based IDS
۱۱ – Accuracy
۱۲ – Detection Rate
۱۳ – False Positive Rate
۱۴ – Dropout
۱۵ – Recurrent Neural Networks (RNN)
۱۶ – Distributed Denial of Service (DDoS)
۱۷ – Convolutional Neural Networks (CNN)
۱۸ – Deep Belief Networks (DBN)
۱۹ – Hybrid Neural Networks
۲۰ – Transfer Learning
۲۱ – Federated learning
۲۲ – Convolutional Layers
۲۳ – Max Pooling Layers
۲۴ – Long Short-Term Memory (LSTM)
۲۵ – Gated Recurrent Unit (GRU)
۲۶ – Fully Connected
۲۷ – Denial of Service (DoS)
۲۸ – Remote to Local (R2L)
۲۹ – User to Root (U2R)
۳۰ – Probe