Deep Learning Approach for Intrusion Detection System

Niharika A P[1], Thaseen Bhashith D[2], Nithya K G[3],Bhindu T N G[4], Sahana A R[5]

Department of CSE JNN College of Engineering,

***Abstract*- The rapid growth of the Internet and communications has resulted in a huge increase in transmitted data. These data are coveted by attackers and they continuously create novel attacks to steal or corrupt these data. The growth of these attacks is an issue for the security of our systems and represents one of the biggest challenges for intrusion detection. An intrusion detection system (IDS) is tool that helps to detect intrusions by inspecting the network traffic. A system called an intrusion detection system (IDS) observes network traffic for malicious transactions and sends immediate alerts when it is observed. It is software that checks a network or system for malicious activities or policy violations. Each illegal activity or violation is often recorded and notified to an administrator. IDS monitors a network or system for malicious activity and protects a computer network from unauthorized access from users, including perhaps insiders. The intrusion detector learning task is to build a predictive model capable of distinguishing between 'malicious connections' and 'genuine connections'.**

***Keywords*: Cyber security, intrusion detection, malware, machine learning, deep learning, deep neural networks, CNN,**
***Dataset:* Keras**

1. INTRODUCTION

A system called an intrusion detection system (IDS) observes network traffic for malicious transactions and sends immediate alerts when it is observed. It is software that checks a network or system for malicious activities or policy violations. Each illegal activity or violation is often recorded and notified to an administrator. IDS monitors a network or system for malicious activity and protects a computer network from unauthorized access from users, including perhaps insiders. The intrusion detector learning task is to build a predictive model capable of distinguishing between 'malicious connections' and 'genuine connections'. Information and communications technology (ICT) systems and networks handle various sensitive user data that are prone by various attacks from both internal and external intruders. Malicious cyber attacks pose serious security issues that demand the need for a novel, flexible and more reliable intrusion detection system (IDS). An IDS is a proactive intrusion detection tool used to detect and classify intrusions, attacks, or violations of the security policies automatically at network -level and host level infrastructure in a timely manner.

2. LITERATURESURVEY

In [1] The approach involves a deep learning model comprising a Convolutional Neural Network (CNN) with a regularized multilayer perceptron instead of the traditional fully connected feed-forward neural network (FNN). CNN uses convolution as a mathematical operation instead of multiplication or dot product. The experimentation employs the UNSW-NB15 dataset, chosen for its representation of real world network traffic and common vulnerabilities.

In [2] The model uses Keras library as a prototype working on top of the TensorFlow framework. Similarly, the framework offers comprehensive and flexible tools and libraries that support deep learning architectures such as CNN and RNN while enabling seamless exhibition for CPU, GPU, and TPU usage. In [3] The Google Colab serves as a free cloud-based Jupyter notebook environment that supports training machine learning or deep learning models using their computing units. The advantage of this work is that the designed architecture and hyper parameter strategy led to substantial model performance, achieving a high accuracy of 94.4% on the testing dataset. In[4] It employed a learning rate of 0.001. Using deep learning models like CNNs often requires significant computational power, which might be a limitation for some setups or environments. Training such models on powerful GPUs or TPUs, as done in this approach, might limit deployment in resource constrained environments. In [5] The UNSW-NB15 dataset was used for its real world representation, it might not encompass all possible network intrusion scenarios, potentially limiting the model's ability to detect new or uncommon threats.

In [6] Network Intrusion Detection Systems (NIDSs) are essential tools for the network system administrators to detect various security breaches inside an organization's network. An NIDS monitors and analyses the network traffic entering into or exitin from the network devices of an organization and raises alarms if an intrusion is observed.

In [7] The KDD Cup dataset has been widely used as a benchmark dataset for many years in the evaluation of NIDS.

One of the major drawbacks with the dataset is that it contains an enormous amount of redundant records both in the training and test data. In [8] NSL KDD was proposed to overcome the limitation of KDD Cup dataset. Dataset improved the previous dataset in two ways. First, it eliminated all the redundant records from the training and test data. Second, it various difficulty levels based on the number of learning algorithms that can correctly classify the records partitioned all the records in the KDD Cup dataset into various difficulty levels based on the number of learning algorithms that can correctly classify the records. In [9] The advantage of this work is that NIDSs based on this approach achieved very high accuracy and less false-alarm rates. STL achieved a classification accuracy rate more than 98% for all types of classification. STL primarily utilizes unlabelled data for feature learning, and the supervised information is derived from a separate set of labelled data. This separation may result in limited alignment between the learned features and the specific characteristics of the labelled data, potentially impacting classification accuracy. In [10] In the number and types of network attacks, traditional firewalls and data encryption methods can no longer meet the needs of current network security. As a result, intrusion detection systems have been proposed to deal with network threats. we take the NSLKDD dataset for simulation testing. A technology called ADASYN is used, which is an adaptive oversampling algorithm based on the minority class samples. In [11] The overall architecture of the DLNID model consists of seven parts, which are the input layer, encoder layer, multiple convolutional layer, attention layer, Bi-LSTM layer, fully connected layer, and the output layer. In [12] In the first layer, the model accepts the network traffic data from the dataset. In the encoder layer, the model uses the encoder part of the improved stacked autoencoder that has been trained to perform dimensionality reduction on the data. In the multiple convolutional layer, the model uses multiple convolutional operations to extract features from the downscaled data. In [13] In the attention layer, the model uses the CBAM to redistribute the weights of each channel and assign more important channels with higher weights. The advantage here is that the experimental data in this paper adopt the NSL-KDD

dataset , which is an improved version of the KDD99 dataset that addresses the data redundancy problem present in the KDD99 dataset and is one of the benchmark datasets used to evaluate the performance of IDS. In [14] The model's performance might be influenced by the choice of hardware and software stack. The computational requirements of the proposed model, especially during training, are not discussed. Deep learning models can be computationally expensive, and the passage does not address potential resource constraints. In [15] The spread of internet networks across the world, cyber-crime increased dramatically. A network intrusion detection system (NIDS) plays an important role in network security. The deep neural networks algorithm as described in the paper can be described via three main steps. In [16] First: the topology of the model, which describes the number of layers and neurons for each layer with the connections between them. In[17] Second: the forward propagation with its perceptron classifier and activation function used by the artificial neurons. In [18] Third: the back propagation with loss function and optimizer. The model topology involves , Input layer: it initializes data for the

neural network purposes. The used system is based on 125 nodes as the input layer, which is represented by the features of the pre-processed dataset. In [19] Hidden layers: It is the intermediate layer between the input and output layer and place where all the computation is done. The usedsystem based on two hidden layers with (50 neural nodes) for the first hidden layer and (30 neural nodes) for the second hidden layer. To evaluate the DNN-IDS model, NIDS is implemented according to two ways of classification: 1) binary classification (Normal and attack), 2) multi-class classification (Normal, DoS, R2L, U2R, and Probe). In [20] . In this paper, two models (multi-class and binary classification) have been proposed, to use deep learning techniques for detecting network attacks instead of using machine learning rules or signatures. Through this experimental research of multi- class classification, which had been found in the KDD cup 99 datasets, it is shown that supervised learning models, which are DNN, are capable of detecting and classifying with high accuracy (99.98 %), and this detection executed on network packet analysis and connection parameters without packet payload information. In [21] The accuracy of detected dos attacks was very high reach to 99.99% The proposed system utilizes the KDD CUP 1999 dataset, which is extensive, allowing for comprehensive training and evaluation. DNNs typically require large amounts of labelled data for effective training. This may require powerful hardware and infrastructure, potentially limiting the accessibility of the system to organizations with significant computational resource. In [22] s. IDS systems can be broadly categorized into two groups: Signature-based Intrusion Detection System (SIDS) and Anomaly-based Intrusion Detection System (AIDS). Intrusion detection datasets are DARPA, KDD, NSL-KDD and ADFA-LDDARPA (Defence Advanced Research Projects Agency) datasets, particularly those related to network security and intrusion detection, are often used for research and benchmarking purposes. In [23] KDD CUP 1999 dataset is one of the most widely used datasets for evaluating intrusion detection systems. The dataset contains network traffic data and is often used to test the performance of intrusion detection algorithms. In [24] fake profile in online social networks using Machine Learning, Detection of Fake Twitter accounts with Machine Learning Algorithms and Twitter fake account detection. In [25] Artificial neural network (ANN) approach as the computational model. A feed forward neural network (FFN), which is a kind of counterfeit brain organization, is utilized to pass different framework data starting with one hub on then onto the next through the organization's edges without producing a circle. We utilize the multilayer perceptron (MLP) model, which contains an info layer, at least one secret layers, and a result layer. Paper gives brief about Algorithm that is Support Vector Machine, or SVM, a notable Directed Learning approach, is utilized to settle relapse and characterization issues.

In [26] A Network Intrusion Detection System (NIDS) helps system and network administrators to detect network security breaches in their organizations. Identifying anonymous and new attacks is one of the main challenges in IDSs researches. some overviews about ICS and IDS, Industrial control system (ICS) is a general term that encompasses several types of control systems, including supervisory control and data acquisition (SCADA) systems, distributed control systems (DCS), and other control system configurations such as Programmable Logic Controllers (PLC) often found in the industrial sectors and critical infrastructures.

Authors	Research focus	Remarks
Lirim Ashiku Cihan Dagli [1] , 2021	This paper focused on the effectiveness of deep learning architectures for network security.	all possible network intrusion scenarios, potentially limiting the model's ability to detect new or uncommon threats.
Quamar Niyaz [2], 2018	Network Intrusion Detection system (NIDS) employing self-taught learning. Used NSL – KDD dataset to detect anomalies and demonstrate high accuracy.	drawbacks with the dataset is that it contains an enormous amount of redundant records both in the training and test data. NSL-KDD CUP overcome .
Yan fang Fu [3], 2022	This paper Intruduces DLNTD, a Deep Learning Network and Bi LSTM networks for accurate traffic anomaly detection.	Deep learning models can be computationally expensive, and the passage does not address potential resource constraints.
Mohammed Maithem [4], 2021	The KDD CUP 1999 dataset preprocessing using z-score normalization and one- hot encoder. DNN with ReLU activation and Adam optimization is constructed	utilizes the KDD CUP 1999 dataset, which is extensive, require powerful hardware and infrastructure, potentially limiting the accessibility.
Patrick Vanin [5], 2022	The influence of dataset quality on model performance, CNN convolutional neural networks (CNN) to address larger datasets.	Highlights the main issue for IDS is their complexity and their low accuracy for minor classes, unavailability of the recent dataset
Ansam Khraisat [6], 2019	The evolution of malicious software (malware) poses a critical challenge to the design of intrusion detection systems (IDS).	cybersecurity threats have evolved since then. As a result, the dataset may not accurately reflect contemporary attack methods and strategies
Ahmad Hijazi [7], 2018	overviews about ICS and IDS, Industrial control system (ICS) is a general term that encompasses several types of control systems,	can be computationally intensive, requiring substantial resources for training and inference.

Mrs. S. Radhika [8], 2023	The development of intrusion detection systems (IDS) for identifying and categorising both network-level and host-level cyberattacks frequently makes use of machine learning techniques.	The system heavily relies on the quality and relevance of the dataset. If the KDD dataset does not represent the current threat landscape accurately.
B. Pavan Kumar [9], 2022	Machine learning techniques are being widely used to develop an intrusion detection system (IDS) for detecting and classifying cyberattacks at the network-level and host-level in a timely and automatic manner.	The scalable framework cannot be disclosed due to confidentiality concerns. This lack of transparency may limit the ability of researchers.
Mrs. T. Madhavi Kumari [10],2021	cyber security method, keeps track of the condition of the network's software and hardware. Existing IDSs still confront hurdles in increasing detection accuracy,	This scalability ensures efficient processing of large datasets and facilitates real-time analysis.

Table -1: Summarization of Various Authors

3. CONCLUSION

In Conclusion, we proposed a hybrid intrusion detection alert system using a highly scalable framework on commodity Deep Learning Approach for Intelligent IDS hardware server which has the capability to analyse the network and host-level activities. The framework employed distributed deep learning model with DNNs for handling and analysing very largescale data in real-time. The DNN model was chosen by comprehensively evaluating their performance in comparison to classical machine learning classifiers on various benchmark IDS datasets. In addition, we collected host based and network-based features in real-time and employed the proposed DNN model for detecting attacks and intrusions. In all the cases, we observed that DNNs exceeded in performance when compared to the classical machine learning classifiers. Our proposed architecture can perform better than previously implemented classical machine learning classifiers in both HIDS and NIDS. To the best of our knowledge this is the only framework which has the capability to collect network level and host-level activities in a distributed manner using DNNs to detect attack more accurately.

REFERENCES

[1] Lirim Ashiku Cihan Dagli, "Network Intrusion Detection System using Deep Learning", ScienceDirect, 185, pp.239-247, (2021).

[2] Quamar Niyaz, Weiqing Sun, Ahmad Y Javaid, and Mansoor Alam, "A Deep Learning Approach for Network Intrusion Detection System",Neurocomput., vol. 122, pp. 13–23, (2013).

[3] Yanfang Fu , Yishuai Du, Zijian Cao, Qiang Li and Wei Xiang, "A Deep Learning Model for Network Intrusion Detection with Imbalanced Data", Electronics, vol.11,pp.287-302,(2022).

[4] Mohammed Maithem, Dr.Ghadaa A. Al-sultany, "Network intrusion detection system using deep neural Networks", ICMAICT 2020 ,vol.13 , pp.1742-6596 ,(2021)

[5] Patrick Vanin, Thomas Newe, Lubna Luxmi Dhirani, Eoin O'Connell, Donna O'Shea, Brian Lee and Muzaffar Rao, "A Study of Network Intrusion Detection Systems Using Artificial Intelligence/Machine Learning", applied sciences, vol 12, (2022).

[6] Ansam Khraisat, Iqbal Gondal, Peter Vamplew and Joarder Kamruzzaman, "Survey of intrusion detection

systems: techniques, datasets and challenges", Khraisat et al. Cybersecurity, pp.2-20, (2019) PRATEEK

[7] SHRIVASTAVA and RK YADAV ,"Deep Learning Approach for Intelligent Intrusion Detection System", SSRN, (2023).

[8] Ahmad HIJAZI, EL Abed EL SAFADI, Jean-Marie FLAUS, "A Deep Learning Approach for Intrusion Detection System in Industry Network", CEUR-WS.org, vol2343,(2021)

[9] Mrs. S. Radhika, K. Navyasree, V. Priyanka, P. Somya, "Deep Learning Approach for Intelligent Intrusion Detection System", Eur. Chem. Bull, vol 12(S3), pp.(2174 – 2178),(2023)

[10] B. Pavan Kumar , J. Asha Jyothi , A. Mounika, "Deep Learning Approaches for Intelligent Intrusion Detection System" , Journal of Cardiovascular Disease Research,VOL13, pp.:0975-3583, 0976-2833, (2022).

[11] Mrs. T. Madhavi Kumari, Aziz Ullah Karimy, "Intelligent Intrusion Detection System Using Deep Learning and Extreme Machine Learning Algorithms", International Journal of Creative Research Thoughts (IJCRT.org), Volume 9, Issue 11 November 2021.

[12] Mr. Venkatram Vennam, Mohammad Abdul Bari Qureshi, Md Amer, Mohammed Abrar Ahmed, Mohd Anas Tayyeb ,"Intelligent Intrusion Detection System Using Deep Learning" , International Journal of Mechanical Engineering, vol. 7 No. 6 June, 2022.

[13]] Mr. Venkatram Vennam, Mohammad Abdul Bari Qureshi, Md Amer, Mohammed Abrar Ahmed, Mohd Anas Tayyeb ,"Intelligent Intrusion Detection System Using Deep Learning" , International Journal of Mechanical Engineering, vol. 7 No. 6 June, 2022.

[14] Kwangjo KIM, "Intrusion Detection System Using Deep Learning and Its Application to Wi-Fi Network", IEICE TRANS. INF. & SYST., VOL.E103–D, NO.7 JULY 2020.

[15] JAN LANSKY,SAQIB ALI , MOKHTAR MOHAMMADI , MOHAMMED KAMAL MAJEED,SARKHEL H. TAHER KARIM,SHIMA RASHIDI, MEHDI HOSSEINZADEH , AND AMIR MASOUD RAHMANI, "Deep Learning-Based Intrusion Detection Systems: A Systematic Review", IEEE Access,vol 9,pp.101574- 101599,(2021).