DEEP LEARNING APPROACH FOR INTELLIGENT INTRUSION DETECTION SYSTEM

Mrs. S. Radhika[1], K. Navyasree[2], V. Priyanka[3], P. Somya[4]

Article History: Received: 08.02.2023 **Revised:** 23.03.2023 **Accepted:** 08.05.2023

Abstract

With more people using the internet, there are more instances of cyberattacks, when a person may be subjected to threats, extortion, or harassment. The attack could take the form of psychological pressure or the theft of a person's password.
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. Due to the lack of a thorough evaluation of the effectiveness of numerous machine learning algorithms in the current approaches.
A kind of deep learning model called Deep Neural Network (DNN) is being researched to find and categorise unanticipated and unplanned cyberattacks in order to develop an effective IDS. This kind of research makes it easier to choose the optimal algorithm for reliably identifying upcoming cyberattacks. Extensive experimental testing has shown that DNNs perform better than conventional machine learning classifiers. This paper suggests a framework for hybrid DNNs that is very scalable and can track network traffic in real time.

[1]Assistant Professor, Sridevi Women's Engineering College, Computer Science and Engineering, Sridevi Women's Engineering College, B. Tech IVYear, Hyderabad, India
[2]Computer Science and Engineering, Sridevi Women's Engineering College, B. Tech IVYear, Hyderabad, India
3Computer Science and Engineering, Sridevi Women's Engineering College, B. Tech IVYear, Hyderabad, India

Email: [1]radhikasefur@gmail.com, [2]navyasri3042@gmail.com, [3]somyaperabathula@gmail.com, [4]priyankavedurumudi@gmail.com

DOI: 10.31838/ecb/2023.12.s3.271

1. Introduction

Spiteful cyberattacks pose serious security difficulties that call for the creation of a cutting-edge, adjustable, and more reliable intrusion detection system. (IDS). Proactive intrusion detection systems, or IDSs, are used to quickly and automatically spot intrusions, attacks, or security policy violations at host- and network-level infrastructure. Two types of intrusion detection that are based on invasive actions are network-based intrusion detection systems (NIDS) and host-based intrusion detection systems. (HIDS). NIDS stands for a network behaviour-based IDS system. Switches, routers, and network taps use network equipment to duplicate network behavioural patterns, which are then evaluated to look for assaults and other potential dangers buried in network data. A HIDS is an IDS that searches for attacks by analysing system activity as it is captured in multiple log files running on the local host machine. The log files are collected using nearby sensors. While HIDS relies on data from log files, such as those for each system's sensor logs, system logs, software logs, file systems, disc resources, user account data, and others, NIDS examines each packet's contents as it travels through network traffic flows. Many businesses employ a mix of NIDS and HIDS. Utilizing misuse detection, anomaly detection, and stateful protocol analysis, network traffic flows are examined. In order to discover the assaults, misuse detection employs filters and pre-set signatures. Human input is used to continuously update the signature database. The unexplained assaults cannot be found using this method at all, but the known attacks can be accurately located using it. Anomaly detection employs heuristic techniques to find unrecognised harmful activity

2. Existing System

Since harmful attacks are evolving constantly and are occurring in extremely high rates, there are numerous difficulties that must be addressed. For further investigation by the digital security industry, many malware datasets are publicly available. The performance of different machine learning algorithms on diverse freely accessible data sets hasn't been thoroughly examined in any existing study, nevertheless. The publicly accessible malware datasets need to be regularly updated and benchmarked since malware is dynamic and its attacking strategies are always evolving.

2.1 Disadvantages

- When threatened, blackmailed, or harassed, malicious cyberattacks pose major security risks.
- The entire detection procedure is slowed down by this huge data, which also increases the likelihood of erroneous results.

Problem Statement

A rising number of people are using the internet. It did draw malicious individuals, though, who want to harm a computer network. An assault can be recognised in the early phases, but by the fourth or fifth step, the system has been totally penetrated. The performance can be enhanced even more by employing cutting-edge techniques and additions.

Proposed statement

The use of a scalable framework on a server with commodity hardware to run a hybrid intrusion detection alert system that can assess host- and network-level activity is suggested. For processing and real-time analysis of extremely huge amounts of data, the structure used deep learning model using DNNs. The DNN model was selected after thorough comparison of their performance to that of convolutional machine learning algorithms on multiple IDS datasets. In order to use the suggested DNN model for identifying assaults and intrusions, we also continuously gathered host- and network-based information. In all 8 instances, we found that DNNs performed better than convolutional machine learning algorithms. In both host based and network based our suggested architecture can outperform classical machine learning classifiers currently in use. As far as we are aware, this is the only framework that can distribute gather host- and network-level actions in order to more effectively detect attacks.

Advantages

- This can successfully work in identifying unplanned and unexpected cyber-attacks at host level and network level.
- Deep learning model can absorb complex vast datasets and may offer correct results in less time.

System design:

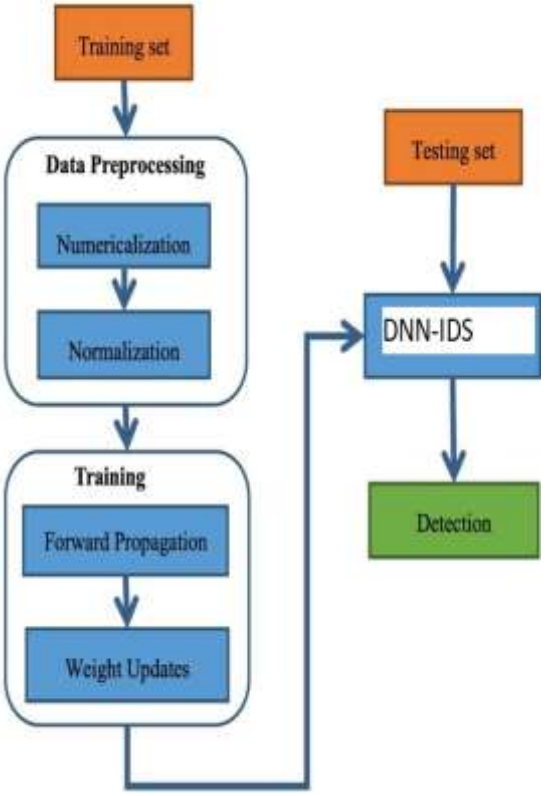

Figure no 1: Architectural view for Deep Neural Network

Modules

Four modules are used to run this project

- Upload Dataset ()
- Pre-process Dataset ()
- Training module ()
- Machine learning module ()

Module Descripton:

- **Upload Dataset module**: Upload Dataset module is used to upload KDD dataset. In the study of intrusion detection methods, the KDD data set serves as a benchmark. In this technique detection framework, data collecting is the primary step that must be completed initially. Data gathering is carried out for the benefit of the machine learning, which generates future predictions.
- **Preprocess dataset**: In this dataset we convert alphabet values to numerical values as 0 and 1 and remove inconsistent data or the incorrect data. A real-world data may contain noisy data, outliers, missing values that cannot be directly used for algorithms in machine learning.
  In this project, the data set is first preprocessed and the data is leaved by removing these unnecessary data which make the predictions accurate.
- **Training module**: In training model we evaluate the performance with various machine learning algorithms.
- **Machine Learning module**: Using this module we train dataset with multiple machine learning algorithms and evaluate their performance and whatever algorithm performing well will use that algorithm to predict any threat.

Implementation

In the implementation first download python version 3.7 along with some functions which are needed for the project. After uploading NSL KDD data set, preprocess the data set to allocate numeric values to each attack names.

Then click on SVM, Random Forest and DNN algorithm to get its prediction accuracy out of which DNN yields better accuracy when compared to other classical machine learning classifiers. Here are some of the screens regarding the implementation.

Figure no 1.1 : Upload NSL KDD Dataset

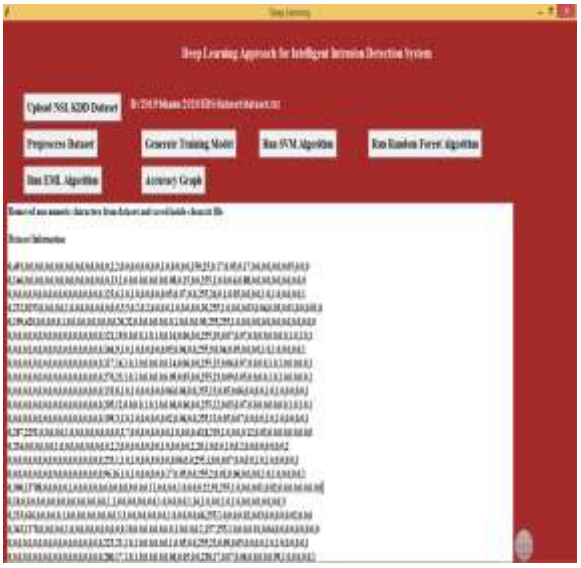

Figure no 1.2: Pre-process dataset

Figure no 1.3: Generate training model

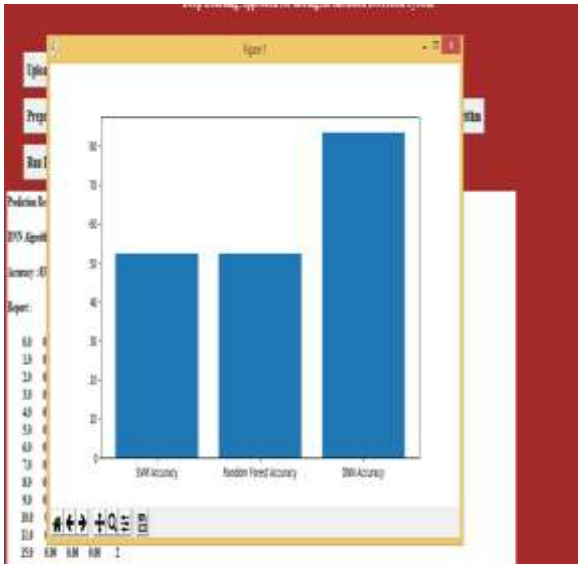

Figure no 1.4: Accuracy Graph

Future scope

In future, further research should consider other Deep learning algorithms to improve the performance in detection accuracy and time complexity and it should not only prevent at the host level, but also within networks of organizations.

By including a module for tracking DNS and BGP events in the networks, the performance of the suggested

framework can be further improved. By adding extra nodes to the current cluster, the suggested system's execution time can be sped up.

3. Conclusion

Deep learning algorithms are highly effective in classifying and developing and intrusion detection system for detecting different types of attacks.

The main objective of using Deep learning methods is it performs better and deals with complex datasets when compared to classical machine learning algorithms and also detect threats with high accuracy. This framework is the only one with the capacity to distribute gather host- and network-level activities using DNNs to more precisely detect attacks.

4. References

Levitt, K. N., Heberlein, L. T., & Mukherjee, B. (1994). monitoring for network intrusions. IEEE Network, 8(3), pp. 26–41.

A. Larson, D. (2016). Attacks that cause a widespread denial of service are preventing a flood. 5-7 in Network Security, 2016(3).

Staudemeyer, R. (2015). Using recurrent neural networks for intrusion detection with extended short-term memory. 136–154 in South African Computer Journal 56(1).