International Journal of Innovative Research in Science and Engineering
Volume 9, Issue 04, 2023
www.ijirse.com
ISSN: 2454-9665

COMPARATIVE ANALYSIS ON DEEP LEARNING METHODS FOR BOT DETECTION

H. Lakshmi Sathvika, Ch. Vamsi Krishna, D.Rohith Kumar, B. Chandra Lokesh

*(Under the guidance of Dr. **R. Lalu Naik**, Professor, Department of Computer Science and Engineering, Tirumala Engineering College)*

ABSTRACT

The Internet Era has been developing for the past decade or so. Lately, in the name of internet exploration, a lot of data transfer has been happening between parties world-wide. The data transfer includes both personal and public data majorly through websites or applications. There is also a lot of increase in demand for Internet of things (IoT) devices. Along with the rise in the use of IoT devices, cyber-attacks are also increasing on these devices. Among them, botnet-based attacks are very high according to studies. Cyber attackers are taking advantage of the vulnerabilities in these systems or websites. With these kinds of intrusions, there are a lot of disasters that happened in the past. There is a necessity to prevent or at least control such attacks.

In order to decrease the number of attacks, a system to identify such harmful attempts has become crucial. There are various types of botnet attacks and various methods to identify them. In this project, we will use different kinds of Deep Learning algorithms to identify Botnet attacks on IoT datasets and compare the results obtained from those algorithms. We will be using three different datasets to work in this project. Deep learning models are implemented in MATLAB 2022b version.

Keywords— IoT, Bot, Deep Learning, Security.

I. INTRODUCTION

The proliferation of IoT devices has made life incredibly convenient and trouble-free. A device that can communicate and send data via the internet when combined with hardware such as sensors, controllers, and software is known as a "internet of things" device. Thanks to a number of applications, such as smart agriculture, smart offices, and smart homes, these devices are becoming a crucial part of our daily life.

The Internet of Things offers all of these advantages,but at the expense of security.There are many instances of cyber attacks due to the accessibility of data over the internet. The industrial sector has occasionally been the subject of cyberattacks, which significantly damaged finances and data.

II. LITERATURE SURVEY

IoT devices are vulnerable to attacks in 57% of cases, with attacks ranging in severity from mild to severe, according to a study on the use of IoT devices in the United States between 2018 and 2019; 41% of attackers take use of such device vulnerabilities [1]. Industries utilise complex methods to defend against these attacks, such as real-time network monitoring and rigid network modelling. However, the majority of these

cyberattacks involve bots, which are automated malicious programmes that may quickly outperform subpar approaches. Everyone who uses a computer at any time could fall subject to a cyberattack. Attacks come in many different forms,from phishing to password cracking[2].

III. PROPOSED SYSTEM:

System Bot detection using Deep Learning Methods mainly focus on detection of Bot based attacks on IOT devices. It also works on large datasets. Here we use neural network for bot detection. Our motive is to apply various deep learning models like pattern recognition networks, feed forward networks, and cascade feed forward networks for classifying and detecting bot attacks. We will be doing feature extraction using correlation, to get the most related feature to classify and identify the attacks. It also works not only on social media but also on other platform.

IV. DATASET AND MODEL

We discovered an ML-based NIDS dataset that was provided by the Australian University of Queensland. They are offering two datasets with different feature sizing. Only 8 basic NetFlow features are used to produce Version 1 datasets, compared to 43 expanded Net Flow features in Version 2 datasets. We only used version 1 of the datasets because version 2 was incompatible with our system. Datasets NF-UNSW-NB15, NF- ToN-IoT, NF-BoT-IoT, and NF-CSE-CIC-IDS2018 are being made available by them [7]. A collection of the detailed feature descriptions may be found in Table 3.1.

In order to determine the optimum model with the highest efficiency, we intend to examine numerous models that are already in the literature. The implementation of a fundamental pattern recognition network, a feed-forward network, and a cascade neural network became our focus. The training function employed distinguishes a pattern network from a feed-forward network. For training a feed-forward network, the levenberg-Marquardt algorithm is employed, however when training a pattern network, scaled conjugate gradient training function is utilised. The following is a list of model parameters:

- Data Division: Random(70% training, 15% validation, 15% testing)
- Number of hidden layers: 2

ISSN: 2454-9665

(a) NF-BoT-IoT

Type	Count
Benign	13859
Reconnaissance	470655
DDoS	56844
DoS	56833
Theft	1909

Type	Count
Benign	270279
DDoS	326345
DoS	17717
Scanning	21467
Injection	468539
Password	156299
XSS	99944
Backdoor	17247
mitm	1295
Ransomware	142

- Number of output features: 2
- Number of neurons in hidden layers: 10
- Training functions: scaled conjugate gradient in patternnet, levenberg-marquardtin feed-forward and cascade
- Max number of epochs: 1000
- Minimum validation checks: 6
- Loss function used: Cross-entropy in pattern net, Mean squared error in FFNN,andCFNN

(b) NF-ToN-IoT

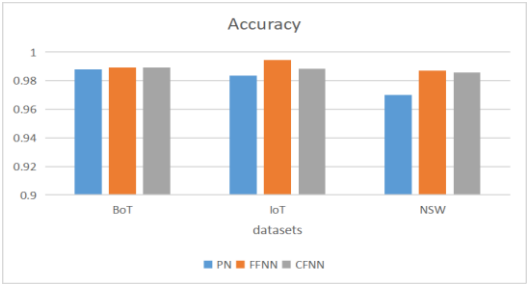
Type	Count
Benign	1550712
Exploits	24736
Fuzzers	19463
Reconnaissance	12291
Generic	5570
DoS	5051
Analysis	1995
Backdoor	1782
Shellcode	1365
Worms	153

(c) NF-UNSW-NB15

Model-Dataset	NF-BoT-IoT	NF-ToN-IoT	NF-UNSW-NB15
PN	0.987817	0.983768	0.970026
FFNN	0.989243	0.994399	0.986918
CFFN	0.989172	0.988216	0.985824

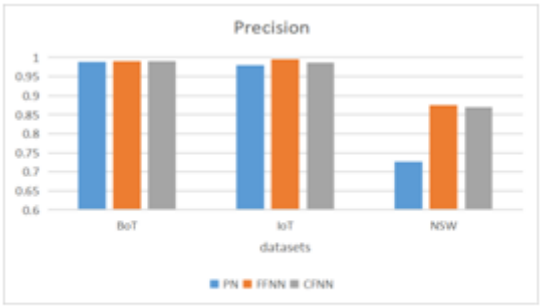
V. RESULTS

We can see from the images that the pattern net is not particularly promising as the amount of the dataset grows. In almost all datasets, FFNN and CFFNN perform pretty similarly, however when we ran the cascade and feed forward by increasing the number of hidden layers (i.e., to 3), cascade outperformed feed forward. With a dataset larger than the three tested, this can be investigated further.

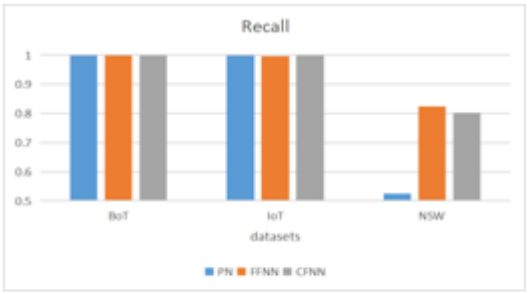


(a) Accuracy of each model of datasets

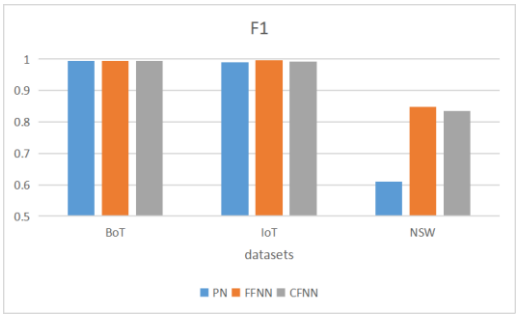
ISSN: 2454-9665



(b) Precision of each model of datasets



(c) Recall of each model of datasets



CONCLUSION

In summary, we can say that feed-forward and cascade neural networks are working almost similarly less or more on all datasets, whereas pattern net is not very promising as the dataset size increases. This can be clearly observed in the pictures shown in Figure 4.1. We did the pre-processing in python and started to implement the deep- learning model as well in python using PyTorch, but then found out that Matlab provides more inbuilt functions in building models so shifted to Matlab for better deep-learning implementation. There is only a very

ISSN: 2454-9665

slight difference in the accuracies of feed-forwardand pattern net even with the increase in dataset size which is very good, but the sincethe model pattern is somewhat different in cascade it takes a lot more time to run.

REFERENCES

[1] "Research about cyber attacks." [Online]. Available: https://unit42.paloaltonetworks.com/iot-threat-report-2020/

[2] "Different types of attacks." [Online]. Available: https://www.csoonline.com/ article/3295877/what-is-malware-viruses-worms-trojans-and-beyond.html

[3] "H. Alkahtani and T. H. Aldhyani, "Botnet attack detection by using cnn-lstm modelfor internet of things applications," *Security and Communication Networks*, vol. 2021, 2021.

[4] A. Derhab, A. Aldweesh, A. Z. Emam, and F. A. Khan, "Intrusion detection system for internet of things based on temporal convolution neural network and efficient feature engineering," *Wireless Communications and Mobile Computing*, vol. 2020,2020.

[5] "Details about bot and botnets."[online]Available: https://www.netscout.com/what-is/bot

[6] "Bot net." [online]. Available: https://datadome.co//learning-center/how-to-detect-mitigate-botnets/

[7] "Datasets." [online] Available: https://staff.itee.uq.edu.au/marius/NIDS_datasets/#