International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)
doi : https://doi.org/10.32628/IJSRST52310217

Detection of Cyber Attacks in Network Using ML

Dr. J Siva Prashanth[1], Dhondi Deepak[2], Bandi Sai Teja[2], Damera Anil[2]
[1]Assistant Professor, Computer Science and Engineering, Anurag Group of Institutions, Hyderabad, Telangana, India.
[2]B. Tech Student, Computer Science and Engineering, Anurag Group of Institutions, Hyderabad, Telangana, India.

ARTICLEINFO

Article History:
Accepted: 01 March 2023
Published: 12 March 2023

Publication Issue
Volume 10, Issue 2
March-April-2023

Page Number
179-183

ABSTRACT

Cybercrime is on the rise everywhere and takes advantage of different flaws in the computing environment. Paying ethical hackers boosted their focus on finding flaws and suggesting solutions. Due to machine learning's success in solving problems related to cyber security, it has recently become a topic of significant relevance. Major concerns in cyber security, such as intrusion detection, malware categorization, and detection, have been addressed using machine learning approaches. Although though it cannot fully automate a cyber security system, machine learning may be able to identify cyber security threats more effectively than other software-oriented approaches, which lessens the stress on security analysts. In this research, we suggest using a machine learning model to identify the network attack. To obtain reliable predictions, some machine learning approaches, such as Random Forest, SVM, Gradient Boosting have been applied. The dataset CSE-CIC-IDS2018 was used to train the model. As a result, efficient adaptive techniques, including different machine learning algorithms, can increase detection rates. The basic objective is to ascertain whether the network is being attacked. Random Forest gave highest accuracy of 99.99%.

Keywords: Machine Learning, Random Forest, SVM, Cyber Attack.

I. INTRODUCTION

Nowadays, computer networks are quite crucial. There are several internet-based services that are used daily, including voice over IP, banking, file sharing, online games, social media, and others. Even though, there are now far more harmful operations taking place on the internet [1]. According to McAfee, "ransomware assaults," a category of malware that aims to prevent a user from using their computer until a predetermined sum of money is paid, have climbed by 118% in 2019[2]. To defend networks from such attacks, dozens of behavior-based detection approaches have been proposed. Using machine learning techniques, the primary issue with these techniques is lowering false alarms. [3–14]. Today's infrastructure for automatically learning features from raw data is provided by machine learning. This

benefit enables the scientists to apply machine learning techniques in several fields, such as computer networks, picture and voice recognition, and natural language processing.

II. LITERATURE SURVEY

Khraisat, and team proposed a paper. This paper presents that, it is getting harder to correctly identify breaches as cyberattacks become more sophisticated. If the attacks are not prevented, security features like data confidentiality, integrity, and availability will become worthless. To combat threats to computer security, several intrusion detection techniques have been developed in the literature. A description of current IDS, a thorough review of significant recent studies, and a list of the datasets typically used for evaluation are all included in this survey study. In order to make computer systems more safe, it also outlines evasion tactics that attackers employ to evade detection and analyses upcoming research challenges to address them. [3]

O. Elejla, and team have proposed an article and made this suggestion. Attacks are a challenging and serious issue on today's Internet, causing financial harm to both individuals and companies. The most frequent attacks against Internet Protocol version 6 are DoS and DDoS assaults employing ICMPv6 messages. They are widespread because the ICMPv6 protocol is required for every IPv6 network to function properly. The new features of IPv6 prevent IPv4 intrusion detection systems (IDSs) from addressing its security issues, such as ICMPv6-based attacks. Nonetheless, IPv4 IDSs can operate in an IPv6 environment. As a result, several IDS have either been developed specially to detect IPv6 threats or have been expanded from existing IPv4 to support IPv6.[4]

M. Idhammad, and his team proposed article for attack detection. This article explains how cutting-edge Machine Learning (ML) methodologies have been used to detect attacks, which continue to pose a danger to the Internet. Unsupervised ML methods, on the other hand, identify attacks by examining incoming network traffic. Large amounts of network traffic data, poor detection precision, and high false positive rates present challenges for both systems. This paper provides a network entropy estimate, clustering, and trees algorithm-based online serial semi-supervised ML technique for DDoS detection. [6]

F. Hossain and team have proposed article on a trustworthy Cyber-attack detection model, is a model that serves as a safety net for user of contemporary technical equipment and a helper for network administrators. By examining network data patterns, the article seeks to advance a CADM for classifying cyberattacks. CADM uses the ensemble classification method to determine attack-level detection accuracy. Important features have been extracted using LASSO. It offers greater visualization capabilities and can handle huge datasets. The classification of network traffic data has been done using an ensemble method that combines the gradient boosting and random forest methods. The Random Forest approach trains each tree in parallel while the Gradient Boosting technique builds weak learning. [21]

Sandosh and his team have proposed a paper for detection of intrusion. In this paper presents, the quick development of cloud computing technology has made it possible for devices across a wide range to connect with ease. This implies a pool of resources that is shared, allowing users to access the data from anywhere in the world. Such a structure faces cyber-security-related difficulties that make it susceptible to outside attacks. To protect the system from intrusions and attacks, an IDS is necessary. The current IDS, however, are unable to effectively combine high accuracy with minimal complexity and speed. To reduce unused spaces, preprocessing is done first using outlier identification. Eventually, the modified K-means clustering data segmentation technique is created. To further classify the attacks, K nearest neighbors (KNN) is employed. [22]

III. PROPOSED SYSTEM

The approaches outlined above used machine learning to identify malicious network activity. Although they performed admirably, combining several machine learning models allows one to take use of each model's strengths and identify attacks far more effectively. In this article, we suggest a general architecture to combine and benefit from any machine learning methods. In this paper, we'll show how to tell whether a network is safe or under attack. If under attack, we can identify the attack's type. using machine learning Our dataset is trained using random forest. Here, the CSE-CIC-IDS2018 dataset from the AWS academy is being used. Parameters used from dataset are source IP address, destination IP address, flag, host log files, port numbers, packets, and payload (TCP/IP header parameters). Since parameters used are unique, accuracy will be greater.

Advantages:

- Attack and type of attack is detected.
- Different network parameters are used which results in higher accuracy.
- Speed of execution is faster compared to other projects.

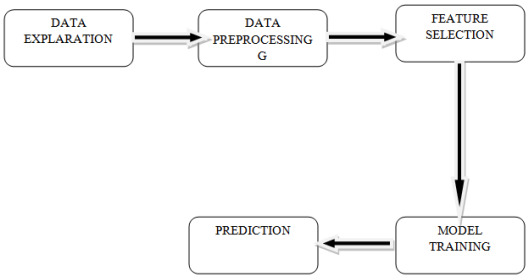

Fig 2.1 System Architecture

IV. IMPLEMENTATION

Random Forest:

The Random Forest calculation is a well-known directed ML method that is frequently utilized in machine learning (ML) for grouping and relapse tasks. We all aware that there are a lot of trees in forest, and that the more trees it has, the stronger it will be.

K Nearest Neighbor (KNN):

A non-parametric, regulated learning classifier, the k-nearest neighbor method, more commonly referred to KNN, makes use of the location to predict or describe the collection of a single piece of information.

Support Vector Machine (SVM):

SVM is a method of directed ML that can be utilized for both characterization & relapse. Even though we also see problems with relapse, the engagement is better. The goal of SVM strategy is to identify hyperplanar information foci in an N-layered space.

Gradient boosting:

In ML, gradient boosting is a type of assistance. Understanding that the best future model, when converged with previous models, reduces the total forecast error is necessary.

V. RESULTS

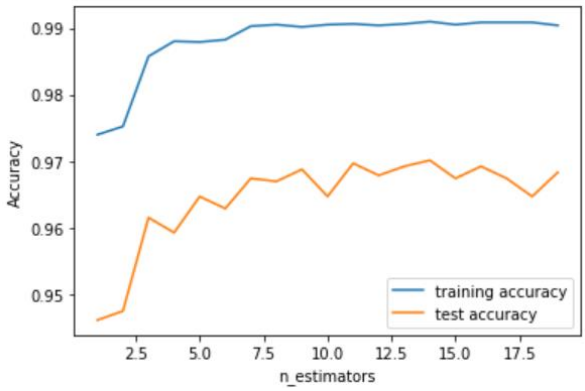

Fig 5.1 Random Forest with accuracy 99.99%

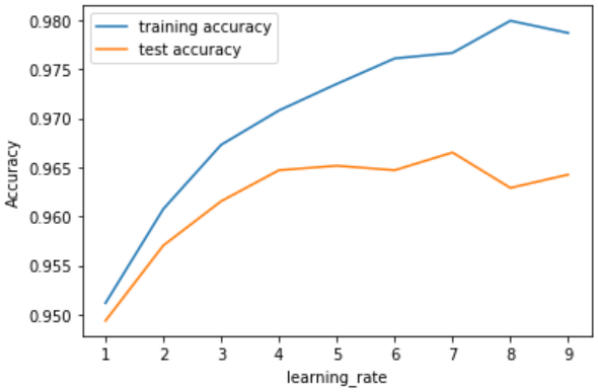

Fig 5.2 Gradient Boosting with accuracy 98.0%

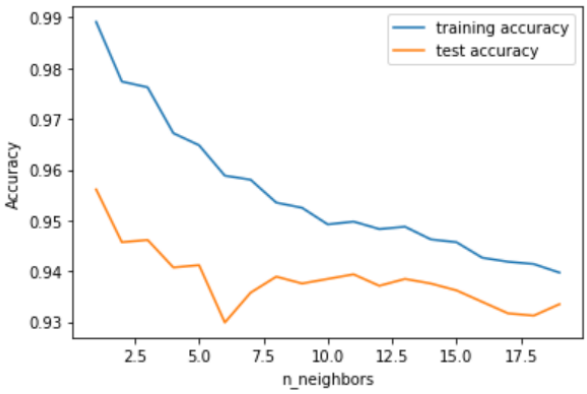

Fig 5.3 KNN with accuracy 99.94%

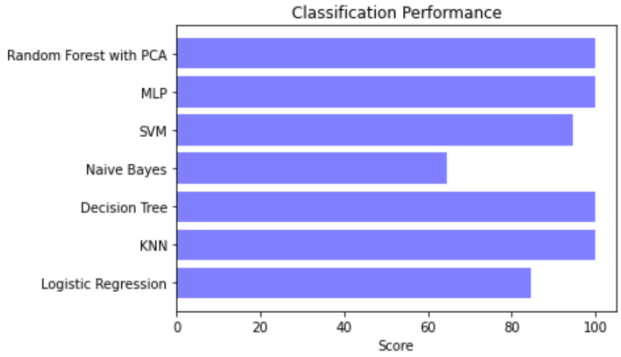

Fig 5.4 Comparison Algorithms vs Accuracy

VI. CONCLUSION

A low-cost machine learning model for cyberattack detection is presented in this paper. According to experimental findings, this study was very effective for all types of machine learning structures with different hyperparameters, interestingly correcting incorrect labels up to 75% of the time. The CSE-CIC-IDS2018 dataset was utilized in this study to test a variety of machine learning methods, with Random Forest providing the greatest accuracy of 99.99%. Other algorithms, like SVM (94.5%) and KNN (99.94%), also performed well.

VII. REFERENCES

[1]. papers/sophoslabs-uncut-2020-threat-report.pdf, 2020.

[2]. https://www.mcafee.com/enterprise/en-us/assets/reports/rp-quarterly-threats-aug-2019.pdf, 2019.

[3]. Khraisat, A., Gondal, I., Vamplew, P. et al. Survey of intrusion detection systems: techniques, datasets and challenges. Cybersecur 2, 20 (2019).

[4]. O. Elejla, B. Belaton, M. Anbar, and A. Alnajjar, Intrusion detection systems of ICMPv6-based DDoS attacks, Neural Computing and Applications, vol. 30, no. 1, pp. 45–56, 2018.

[5]. M. H. Haghighat and J. Li, Edmund: Entropy based attack detection and mitigation engine using netflow Data, in Proc. of 8th International Conference on Communication and Network Security, Chengdu, China, 2018, pp. 1–6.

[6]. M. Idhammad, K. Afdel, and M. Belouch, Semi-supervised machine learning approach for DDoS detection, Applied Intelligence, vol. 48, no. 10, pp. 3193–3208, 2018.

[7]. D. S. Terzi, R. Terzi, and S. Sagiroglu, Big data analytics for network anomaly detection from netflow data, in Proc. of 2017 International Conference on Computer Science and Engineering, Antalya, Turkey, 2017, pp. 592–597.

[8]. J. M. Vidal, A. L. S. Orozco, and L. J. G. Villalba, Adaptive artificial immune networks for mitigating DoS flooding attacks, Swarm and Evolutionary Computation, vol. 38, pp. 94–108, 2018.

[9]. R. Wang, Z. Jia, and L. Ju, An entropy-based distributed DDoS detection mechanism in software-defined networking, in Proc. of 2015 IEEE Trustcom/BigDataSE/ISPA, Helsinki, Finland, 2015, pp. 310–317.

[10]. G. Aceto, D. Ciuonzo, A. Montieri, and A. Pescapé, Multiclassification approaches for classifying mobile app traffic, Journal of Network and Computer Applications, vol. 103, pp. 131–145, 2018.

[11]. Tran, K. N., Alazab, M., & Broadhurst, R. (2013, November). Towards a feature rich model considering predicting spam emails containing

malicious attachments & urls. In 11th Australasian Data Mining Conference, Canberra.

[12]. Alazab, M., & Broadhurst, R. (2015). Spam & criminal activity.

[13]. Alazab, M., Layton, R., Broadhurst, R., & Bouhours, B. (2013, November). Malicious spam emails developments & authorship attribution. In Cybercrime & Trustworthy Computing Workshop (CTC), 2013 Fourth (pp. 58-68). IEEE.

[14]. Broadhurst, R., Grabosky, P., Alazab, M., Bouhours, B., & Chon, S. (2014). An analysis about nature about groups engaged in cybercrime.

[15]. Alazab, M., Venkatraman, S., Watters, P., & Alazab, M. (2011, December). Zero-day malware detection based on supervised learning algorithms about API call signatures. In Proceedings about Ninth Australasian Data Mining Conference-Volume 121 (pp. 171- 182). Australian Computer Society, Inc..

[16]. Vinayakumar, R., Alazab, M., Srinivasan, S., Pham, Q. V., Padannayil, S. K., & Simran, K. (2020). A Visualized Botnet Detection System based Deep Learning considering Internet about Things Networks about Smart Cities. IEEE Transactions on Industry Applications.

[17]. Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., Al-Nemrat, A., & Venkatraman, S. (2019). Deep learning approach considering intelligent intrusion detection system. IEEE Access, 7, 41525-41550.

[18]. Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., & Venkatraman, S. (2019). Robust intelligent malware detection using deep learning. IEEE Access, 7, 46717-46738.

[19]. Vinayakumar, R., Alazab, M., Jolfaei, A., Soman, K. P., & Poornachandran, P. (2019, May). Ransomware triage using deep learning: twitter as a case study. In 2019 Cybersecurity & Cyberforensics Conference (CCC) (pp. 67-73). IEEE.

[20]. Srinivasan, S., Ravi, V., Sowmya, V., Krichen, M., Noureddine, D. B., Anivilla, S., & Kp, S. (2020, March). Deep convolutional neural network based image spam classification. In 2020 6th Conference on Data Science & Machine Learning Applications (CDMA) (pp. 112-117). IEEE.

[21]. F. Hossain, M. Akter and M. N. Uddin, "Cyber Attack Detection Model (CADM) Based on Machine Learning Approach," 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), DHAKA, Bangladesh, 2021, pp. 567-572.

[22]. Sandosh, S., Govindasamy, V. & Akila, G. Enhanced intrusion detection system via agent clustering and classification based on outlier detection. Peer-to-Peer Netw. Appl. 13, 1038–1045 (2020).

Cite this article as :

Dr. J. Siva Prashanth, Dhondi Deepak, Bandi Sai Teja, Damera Anil, "Detection of Cyber Attack in Network Using ML", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 10 Issue 2, pp. 179-183, March-April 2023. Available at doi : https://doi.org/10.32628/IJSRST52310217
Journal URL : https://ijsrst.com/IJSRST52310217