Vol 8, Issue 6, 2025 **Impact Factor: 8.535** DOI: https://doi.org/10.31426/ijamsr.2025.8.6.8433

IJAMSR

International Journal of Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

A Comparative Study of Classification Algorithms for Network Intrusion Detection

Dogiparthy Sravankumar [1]

[1] Research Scholar, Department of Computer Science, Mansarovar Global University, Sehore, M.P., India.

Dr. G. Soma Sekhar [2]

[2] Supervisor, Department of Computer Science, Mansarovar Global University, Sehore, M.P., India.

ABSTRACT

Protecting sophisticated communication networks requires intrusion detection systems (IDS). The main purpose of these systems was to detect certain patterns, signatures, and rule infractions. There have been encouraging developments in the area of network intrusion detection thanks to the application of Machine Learning and Deep Learning techniques in recent years. Such methods are able to distinguish between typical and out-of-the-ordinary patterns. In order to efficiently detect network assaults, this study suggests a machine learning-based intrusion detection model that combines PCA and SVM. Common assessment criteria including F1-score, recall, accuracy, and precision are used to assess the suggested model's performance. The SVM classifier outperforms other algorithms like Decision Tree and Naïve Bayes, according on the testing data.

Keywords: *Intrusion Detection, Accuracy, Precision, Machine learning, Network.*

I. INTRODUCTION

"To use machine learning to identify suspicious attacker activity on a network or system"Unauthorised access to computer systems is one of the biggest threats to computer or network security in the modern world. New forms of network assaults are appearing often due to the exponential growth of network applications. We need to change our system to handle suspicious behaviour. It is possible to alert the director's strator the moment an assault or suspicious activity is noticed. In order to identify and redirect assaults, access control systems (IDS) use a host-or network-based approach. Whatever the case may be, attack signatures on these items frequently point to malevolent or suspect intentions. These patterns in network traffic are retrieved from the network when the ID is validated.

IJAMSR

International Journal of Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

Most of the methods utilised by contemporary IDs aren't up to the task of handling the dynamic and intricate nature of cyberattacks on computer networks. Thus, once again, by making use of adaptive processes, we may achieve high detection rates with little false alarms. Machine learning techniques offer suitable communication and accounting expenses. Your expansion signature number and link speed must be met before the matching pattern can quickly verify if an incoming package has a signature and acquire harmful behaviour. Multiple applications can make advantage of IDS.

There has been a lot of buzz about how to incorporate machine learning techniques into SDN. By developing a strategy and doing thorough exploratory research with the NSL-KDD data set, we were able to achieve high accuracy in our entry discoveries and so resolve the issues with the KDD Cup 99. This experimental investigation included five popular and effective machine learning algorithms: RF, J48, SVM, CART, and Naïve Bayes. After applying the feature selection technique to simplify the features, the number of features in the NSL-KDD database was reduced to 13.

A suggested "Intelligent Access Acquisition System" concept that is dynamic and based on a particular AI approach to acquiring access. With the use of basic data mining techniques, strategies that combine abstract reasoning with neural networks may analyse network data and create a network profile. Confusion, abuse, and host-based acquisitions are all a part of the scheme. The rules can be designed to mirror the usual ways security attacks are defined, with a focus on simplicity and comprehensiveness. Machine learning programs have employed a variety of approaches to solve the issue of feature access selection. To find space characteristics in the major character region, the author utilised principal component analysis (PCA). Then, using the Genetic Algorithm, they selected features that correlate to high eigen values. Each model achieved an average accuracy of 73% after being re-trained with 13 reduced features; the others achieved 98%, 85%, 95%, and 86%. The proposal aimed to use a model of a deep neural network to identify and prevent SDN infiltration.

II. REVIEW OF LITERATURE

Golande, Shashikant et al., (2024) Network security is critical in the modern digital world, and intrusion detection systems (IDS) are an integral part of that process, keeping sensitive information safe from hackers. Traditional intrusion detection systems (IDS), which frequently use signature-based approaches, have a hard time adjusting to new threats, need a lot of computing power, and have high false positive rates. Using machine learning approaches, this study investigates how to build a system that can efficiently identify and classify intrusions into networks. We use datasets like NSL-KDD and UNSW-NB15 to conduct our study. We use a mix of supervised learning techniques, including SVM, Random Forests, and Neural Networks. We also preprocess the data and design features thoroughly. Our models show a significant improvement in detecting capabilities and computational efficiency when evaluated using measures like recall, accuracy, precision, and ROC-AUC. Based on our research, it appears that IDS powered by machine learning can outperform traditional systems in terms of network security, particularly when it comes to lowering false positives and keeping up with new threats. Our study not only lays the groundwork for future advancements in IDS using advanced machine learning techniques, but it also highlights their potential. Protecting sensitive information and ensuring continuous operations in the ever-changing world of cybersecurity requires reliable intrusion detection and classification systems for networks.

IJAMSR

International Journal of Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

Intrusion detection systems (IDS) may be made more efficient and accurate with the help of a new method proposed in this study that makes use of machine learning. Our system is able to detect and categorise unexpected and known threats simultaneously by utilising a mix of supervised and unsupervised learning methods. To make sure our models work as well as possible while producing few false positives, we use sophisticated feature selection techniques. When compared to more conventional intrusion detection systems, our experimental findings show considerable improvements in both detection accuracy and processing speed, as confirmed on benchmark datasets. In addition to fortifying network defences, the suggested solution offers a flexible and extensible architecture to address emerging cybersecurity threats.

Ahmad, Iftikhar et al., (2022) Due to its impact on several communication and security sectors, intrusion detection in computer networks is highly significant. It is difficult to identify intrusions into networks. Furthermore, training state-of-the-art machine learning models to detect network intrusion risks requires a vast quantity of data, making network intrusion detection a tough undertaking. The topic of network intrusion detection has lately seen several suggested methods. The ever-increasing number of new dangers, however, presents significant obstacles that existing systems are ill-equipped to handle. In order to build an intrusion detection system for networks, this article examines several methods. We use the correlation between characteristics to choose the best ones from the dataset. Moreover, we offer a thorough functional and performance-oriented AdaBoost-based method for network intrusion detection that is based on these chosen characteristics. For network anomaly identification, we utilised the more current and extensive UNSW-NB 15 dataset, as opposed to the KDD99 dataset utilised by the majority of prior research. Dataset is a compilation of packets sent and received by hosts in a network. The 49 characteristics make up nine distinct threat types: denial-of-service, fuzzy logic, exploit, worm, shellcode, reconnaissance, generic, and analytical backdoor. For the purpose of this research, we compare SVM with MLP. Lastly, we suggest AdaBoost, which uses the decision tree classifier to distinguish between safe and dangerous actions. We kept an eye on the data flowing over the network and labelled it as either dangerous or harmless. Based on the results of the experiments conducted on the UNSW-NB15 dataset, our suggested technique successfully identifies various types of network intrusions on computer networks with an accuracy of 99.3 percent. Network security applications and research sectors will find the suggested system useful.

Cihan, Şeyma et al., (2021) The amount of assaults on networks has grown substantially in tandem with the advancements in network technology. Every day, there is a growing demand for robust intrusion detection systems to keep networks secure and stable. An intrusion detection system that combines deep learning with more conventional machine learning techniques is suggested in this research. Random Forest, Decision Tree, and Deep Neural Network techniques were used to classify the NSL-KDD dataset in this study. Also, in order to reduce the dataset's dimensionality, the Gini index and CFS (Correlation Based Feature Selection) were used to define variable subsets. After the dataset was reduced to 11 variables using the CFS approach, the study found that the Random Forest algorithm produced the maximum accuracy rate of 99.972%. Furthermore, Deep Neural Network achieved an accuracy rate of 99.64% without the need for feature engineering.

IJAMSR

International Journal of Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

Vanitha, L & Mary, Safish. (2017) A Network Intrusion Anomaly Detection System (NIADS) is a piece of hardware or an app that scans incoming network traffic for malicious data. The correct application of distance and similarity metrics between data flowing across the network is critical for the categorisation of incoming data as anomalous. In intrusion detection, categorisation is necessary for spam filtering and undesirable data entry prevention in incoming network data. Intrusion detection systems categorise input to determine the presence or absence of harmful attacks. This article explores the different classification techniques employed in these systems. The purpose of this research is to aid in the search for and development of the optimal offline NIDS by providing a comparative analysis of the several classification algorithms currently in use.

Patel, Hemant. (2013) Intrusion detection system (IDS) research has shown that there are numerous real-world threats and several IDS systems that can detect them, including application-based, host-based, and network-based IDS. This article presents a framework for an intrusion detection system that uses data mining to examine intrusion data and identify data sets that are susceptible to network assaults. Additionally, we cover the fundamentals of data mining as it pertains to discovering intrusion data within the dataset. Identification within the domain of data mining for intrusion detection requirements. We go over a few of the most popular intrusion detection techniques as well, including decision trees, Naive Bayes, Naive Bayes (CFSGSW), and NBTree adaptive NBTree it.

III. PROPOSED METHODOLOGY

Both labelled and unlabelled data comprising vulnerabilities and threats are accepted as input by the suggested system. Before training and testing, data undergoes preprocessing so that key characteristics may be used for classification and feature separation. The SVM classifier is used to train the classification model. Attack recognition makes use of principal component analysis to decide whether to proceed with intrusion detection or maintain regular data flow after discovering an intrusion in the test data.

Data Set

This experiment makes use of the KDD Cup'99 dataset, with 10% of that dataset serving as a training set and the remaining 90% serving as a testing set for identifying network attacks. All four forms of assault are included in the dataset.

- Denial of Service (DOS)
- Probing
- Root to Local (R2L)
- User to Root (U2R)

Principle Component Analysis (PCA)

There are a lot of characteristics in network data that might not be helpful for identifying the kind of assault. Therefore, such superfluous characteristics are culled using principal component analysis. The dimensionality reduction approach known as principle component analysis reduces huge dimensions to tiny ones. This feature reduction improves performance time and boosts attack detection accuracy.

IJAMSR

International Journal of Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

Support Vector Machine (SVM)

The supervised learning technique known as support vector machine can analyse data, identify patterns, and run regressions. Finding the hyper-plane of the training data set is a common way that support vector machines do multi-class classification. When training a multi-class support vector machine, it is possible to create several classes simultaneously.

Evaluation Metrics

The performance of the classification algorithms is measured using the Accuracy, Precision, Recall and F1-Score metrics.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

Table 1: Performance Comparison of Classification Algorithms

Approaches	Accuracy	Precision	Recall	F1-Score
SVM	0.92	0.89	0.92	0.90
Decision tree	0.87	0.85	0.87	0.86
Naive Bayes	0.81	0.80	0.81	0.80

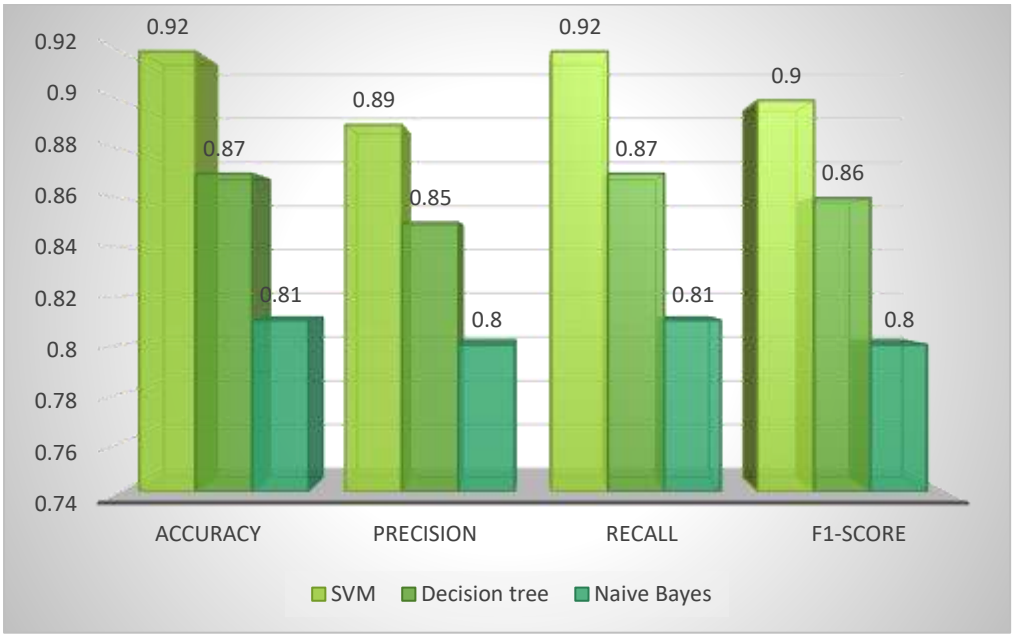

Figure 1: Performance Comparison of Classification Algorithms

The SVM classifier outperforms the other two methods with respect to all metrics measured in Table 1. Achieving an accuracy of 0.92 means that 92% of the network traffic examples are properly classified by the model. With a recall of 0.92 and a precision of 0.89, the model is quite good at detecting real intrusion attempts, and the majority of the predicted attack cases are properly recognised. The F1-score of 0.90 provides additional evidence that SVM performs equally well in terms of recall and precision.

IJAMSR

International Journal of Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

The Decision Tree approach, on the other hand, performs somewhat worse, with an F1-score of 0.86, an accuracy of 0.87, a precision of 0.85, and recall of 0.87. Its detection power is inferior to that of the SVM model, yet it does a respectable job of identifying network activity. However, out of the three methods, the Naïve Bayes classifier performs the worst, with an accuracy of 0.81, precision of 0.80, recall of 0.81, and F1-score of 0.80.

V. CONCLUSION

Nowadays, in the age of high-speed networks, technologies based on machine learning are widely utilised for analysing massive amounts of network traffic data. Achieving a high intrusion detection rate while decreasing false positive and negative states is a critical problem when building intrusion detection systems. This article addressed these concerns by providing an overview of intrusion detection systems (IDS) and an explanation of methods for detecting intrusions in networks using machine learning. Intruder detection using these machine learning based technologies is successful. The development of IDS has been a collaborative effort involving several academics and researchers. To increase the detection rate and decrease the false state, further research and study are needed.

REFERENCES

1. S. Golande, S. Vaidya, A. Pardeshi, V. Katkade, and V. Pawar, "An efficient network intrusion detection and classification system using machine learning," *International Journal of Advanced Research in Science, Communication and Technology*, vol. 4, no. 1, pp. 267–272, 2024.
2. I. Ahmad, Q. Haq, M. Imran, M. Alassafi, and R. AlGhamdi, "An efficient network intrusion detection and classification system," *Mathematics*, vol. 10, no. 3, pp. 1–15, 2022.
3. G. Sajja, M. Jawarneh, P. Ponnusamy, S. Abdufattokhov, M. G. Murugesan, and P. Prabhu, "Machine learning algorithms in intrusion detection and classification," *Annals of the Romanian Society for Cell Biology*, vol. 25, no. 6, pp. 12211–12219, 2021.
4. Ş. Cihan, M. Aydos, and N. Simsek, "A tree based machine learning and deep learning classification for network intrusion detection," *European Journal of Science and Technology*, vol. 31, no. 1, pp. 104–113, 2021.
5. L. Vanitha and S. Mary, "A comparative study of classification algorithms used in network intrusion detection systems (NIDS)," *ARS Journal of Applied Research and Social Sciences*, vol. 3, no. 23, pp. 7–14, 2017.
6. M. Belouch, S. El Hadaj, and M. Idhammad, "A two-stage classifier approach using REPTree algorithm for network intrusion detection," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 6, pp. 389–394, 2017.
7. K. Kumar and J. S. Batth, "Network intrusion detection with feature selection techniques using machine-learning algorithms," *International Journal of Computer Applications*, vol. 150, no. 12, 2016.
8. H. Patel, "Intrusion detection in data mining with classification algorithm," *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 2, no. 7, 2013.

IJAMSR

International Journal of Advanced Multidisciplinary Scientific Research (IJAMSR) ISSN:2581-4281

9. Y. B. Bhavsar and K. C. Waghmare, "Intrusion detection system using data mining technique: Support vector machine," *International Journal of Emerging Technology and Advanced Engineering*, vol. 3, no. 3, pp. 581–586, 2013.
10. A. Mitrokotsa and C. Dimitrakakis, "Intrusion detection in MANET using classification algorithms: The effects of cost and model selection," *Ad Hoc Networks*, vol. 11, no. 1, pp. 226–237, 2013.
11. M. Panda, A. Abraham, S. Das, and M. Patra, "Network intrusion detection system: A machine learning approach," *Intelligent Decision Technologies (IDT) Journal*, vol. 5, no. 4, pp. 347–356, 2011.
12. S. Mukkamala, A. H. Sung, and A. Abraham, "Intrusion detection using an ensemble of intelligent paradigms," *Journal of Network and Computer Applications*, vol. 28, no. 2, pp. 167–182, 2005.