**NETWORK INTRUSION DETECTION SYSTEM BY SUPERVISED MACHINE LEARNING TECHNIQUE**

**ABSTRACT**

Intrusion detection plays an important role in ensuring information security, and the key technology is to accurately identify various attacks in the network. In our study, we explore how to model an intrusion detection system based on supervised machine learning. Moreover, A new supervised machine learning system is developed to classify network traffic data whether it is malicious or benign. We developed a supervised machine learning model that can classify unseen network traffic based on what is learnt from the seen traffic. We used both SVM and ANN learning algorithm to find the best classifier with higher accuracy and success rate.

we study the performance of the model in binary classification and multiclass classification, and the number of neurons and different learning rate impacts on the performance of the proposed model. We compare it with those of J48, Artificial Neural Network, Random Forest and other machine learning methods proposed by previous researchers on the benchmark dataset. The experimental results show that supervised machine learning technique is very suitable for modeling a classification model with better accuracy and that its performance is superior to that of traditional machine learning classification methods in both binary and multiclass classification. The supervised machine learning model improves the accuracy of the intrusion detection and provides a new research method for intrusion detection. The computer security has become a major challenge. Tools and mechanisms have been developed to ensure a level of compliance. These include the Intrusion Detection Systems (IDS). The principle of conventional IDS is to detect attempts to attack a network and to identify abnormal activities and behaviors. The reasons, including the uncertainty in searching for types of attacks and the increasing complexity of advanced cyber-attacks.

**PROBLEM STATEMENT**

* Cyber Security is becoming a severe issue for individuals, enterprises, and governments alike. In a world where everything is on the internet, from cute kitten videos and our travel diaries to our credit card information, ensuring that our data remains safe is one of the biggest challenges of Cyber Security.
* Cyber Security challenges come in many forms, such as ransomware, phishing attacks, malware attacks, and more. India ranks 11th globally in terms of local cyber-attacks and has witnessed 2,299,682 incidents in Q1 of 2020 already.

* The software-based security systems cannot provide promising security to new types of threats.
* The behavior of monitored environment may change after certain period of time and require retraining of the system.
* If the training set itself contains attacks, the system will consider malicious behavior as Normal and false positives and false negatives.

**SOLUTION**

* The goal of a network intrusion detection system is to discover unauthorized access to a computer network by analyzing traffic on the network for signs of malicious activity.
* A network intrusion detection system (NIDS) monitors both inbound and outbound traffic on the network, as well as data traversing between systems within the network.
* Hence, we are using Supervised Machine Learning Algorithm called Support Vector Machine (SVM) and Artificial Neural Network (ANN).
* Support Vector Machine (SVM) In SVM a separating hyper plane defines the classifier depending on the type of problem and available datasets. In case where dataset is one dimensional, the hyper plane is a point, for two-dimensional data it is a separating line, for three-dimensional dataset, it is a plane and if the data dimension is higher, it is a hyper plane. For a linearly separable dataset, the classifier or the decision function will have the form.

**AREA AND DOMAIN**

* The project we are working is under Cyber Security and Machine Learning.

**EXISTING SYSTEM**

* Several real attacks are far less than the number of false alarms raised. This causes real threats to go often unnoticed.
* Noise can severely reduce the capabilities of the IDS by generating a high false-alarm rate.
* Constant software updates are required for signature-based IDS to keep up with the new threats.
* IDS monitor the whole network, so are vulnerable to the same attacks the network’s hosts are. Protocol-based attacks can cause the IDS to fail.
* Network IDS can only detect network anomalies which limit the variety of attacks it can discover.
* Network IDS can create a bottleneck as all the inbound and outbound traffic passes through it.
* Host IDS rely on audit logs, any attack modifying audit logs threaten the integrity of HIDS

**PROPOSED SYSTEM**

* Deep learning algorithms are used since they are more suitable to large-scale data when compared to traditional machine learning algorithms.
* Since it is taking large scale data and the required time is more, we are using medium scale data which takes less time and results are obtained fast compared to the previous one.
* The NSL-KDD dataset is being used to detect network attacks.
* Attacks such as DOS, DDOS, Web Attacks, Bot Net attacks are detected.

**SOFTWARE REQUIREMENTS**

* Language: Python
* OS: Windows 10(64bit)
* Package- Python 3.7
* Software- Anaconda

**HARDWARE REQUIREMENTS**

* Processor: Above 1.5GHZ
* Hard Disk: 80GB
* RAM: 2GB

**Conclusion**

* We have presented different machine learning models using different machine learning algorithms and different feature selection methods to find a best model.
* The analysis of the result shows that the model built using ANN and wrapper feature selection outperformed all other models in classifying network traffic correctly with detection rate approx 94.02%. We believe that these findings will contribute to research further in the domain of building a detection system that can detect known attacks as well as novel attacks.
* In future we can improve this system to detect new attack in the network.