iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

NETWORK INTRUSION DETECTION USING SUPERVISED MACHINE LEARNING TECHNIQUE

M. RAJ KUMAR
Assistant Professor, Department of Computer Science and Engineering, J.B. Institute of Engineering and Technology, Hyderabad.
A. MANI CHANDANA [1]
V. HARINI [2]
UG Students, Department of Computer Science and Engineering, J.B. Institute of Engineering and Technology, Hyderabad.

ABSTRACT
In today's digital era, where most services are delivered through the internet, protecting client and server machines from malicious attacks is essential. Network Intrusion Detection Systems (IDS) play a critical role in identifying and mitigating such threats by analyzing incoming request data to detect potential attack signatures. This paper evaluates the performance of two supervised machine learning algorithms, Support Vector Machine (SVM) and Artificial Neural Networks (ANN), in detecting anomalies within network traffic. The IDS is trained using a comprehensive dataset containing normal and attack signatures. If an attack signature is detected, the request is dropped, and the malicious data is logged for future analysis. Through experimental analysis, we demonstrate that ANN outperforms SVM in terms of accuracy, making it a more reliable choice for intrusion detection. This study highlights the importance of enhancing IDS systems using advanced machine learning techniques to safeguard digital systems against emerging cyber threats.

Keywords:
Networks Intrusion Detection System (IDS), Machine Learning, Support Vector Machine (SVM), Artificial Neural Networks (ANN), Anomaly Detection, Cyber Security, Attack Signatures, Model Performance Evaluation, Request Data Monitoring, Network Security

INTRODUCTION

With the rapid growth of internet usage and the increasing availability of online content, cybercrime has become a significant threat. Intrusion detection is the first step to prevent security attack. Intrusion Detection Systems (IDS) is widely used to detect attacks by analyzing data from various network sources to identify potential security breaches. Network based IDS inspects data packets travelling over the network, using both signature- based and anomaly- based detection techniques. However, anomaly- based IDS faces challenges in identifying novel attacks due to the lack of prior knowledge. To address this, researchers have been exploring machine learning techniques to enhance the capability of IDS in distinguishing between normal and malicious traffic. Despite significant commercial investment and research in the field of intrusion detection since 1980s, anomaly-based IDS has not yet achieved the same level of success as signature-based systems. This project aims to explore the effectiveness of using supervised machine learning techniques, such as SVM and ANN, to improve anomaly- based intrusion detection and enhance network security.

OBJECTIVES

The main objective of this project is to develop an efficient Network Intrusion Detection System by utilizing the Supervised Machine learning Algorithms. In this project, Feature selection techniques like filter and wrapper methods are being used to reduce data dimensionality and improve model performance. By identifying the best performing

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

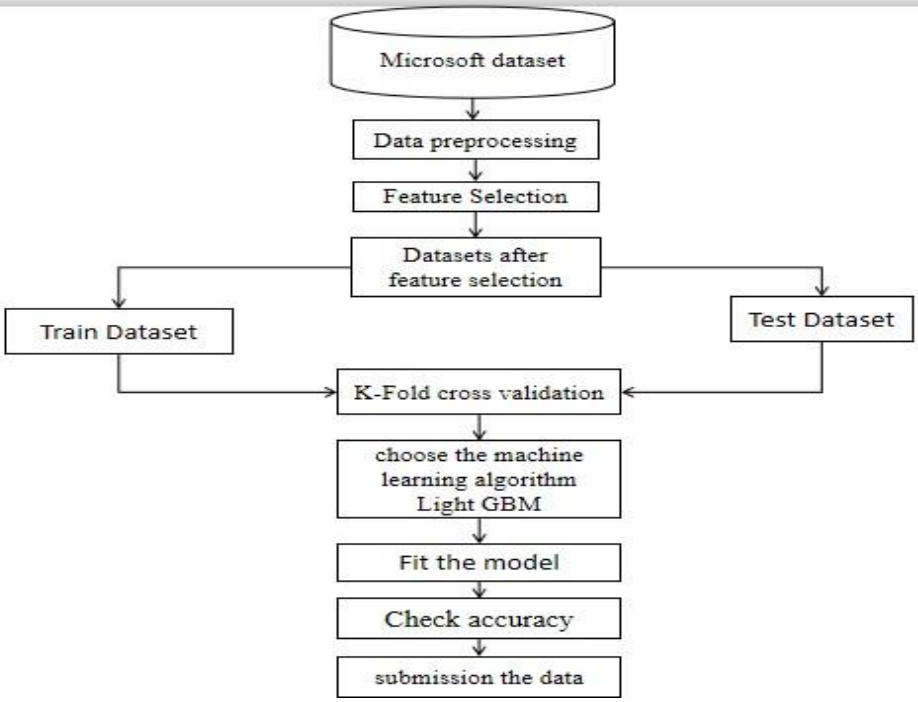
model for Intrusion Detection, we can implement the real-time monitoring and threat prevention. Thus suggesting the future improvements for Intrusion Detection Systems.

METHODOLOGY

A comprehensive project methodology for developing a Network Intrusion Detection System (NIDS) using supervised machine learning can be broken down into distinct stages:

1. **Data Collection** - Data collection primarily involves selecting a suitable, publicly available dataset like NSL-KDD, UNSW-NB15, or CICIDS2017, ensuring it offers diverse attack patterns and relevant network traffic features. Analyzing the dataset's characteristics, including feature types and attack distribution, is crucial for effective preprocessing and model training.
2. **Preprocessing** - Preprocessing for a supervised machine learning NIDS project involves cleaning the network traffic dataset by handling missing values and removing duplicates. Categorical features are transformed into numerical representations using techniques like one-hot encoding, while numerical features are scaled or normalized for consistent ranges. Feature selection or dimensionality reduction methods are applied to enhance model efficiency and reduce noise.
3. **Model Training** - Model training in a supervised learning NIDS project involves feeding the preprocessed network traffic data to selected algorithms like Random Forests, SVM, or Gradient Boosting Machines. The training process adjusts model parameters to learn patterns that distinguish between normal and malicious traffic. Hyperparameter tuning, using techniques like cross-validation, optimizes the model's performance on the training data. This results in a trained model capable of classifying unseen network traffic.
4. **Evaluation** - Project evaluation for a supervised learning NIDS involves assessing the trained model's performance using metrics like accuracy, precision, recall, F1-score, and AUC. A confusion matrix is used to analyze classification errors. Performance is also evaluated across various attack types, and the model's generalization is tested on unseen data to ensure robustness.
5. **Deployment** - Deployment of a supervised learning NIDS involves integrating the trained model into a network environment for real-time traffic analysis. This requires efficient processing of network packets, feature extraction, and rapid classification to generate alerts. Scalability and continuous monitoring are crucial for handling high traffic volumes and adapting to evolving threats.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/



RESULTS AND DISCUSSION

To assess the effectiveness of different models in classifying network traffic, we trained and tested two primary machine learning models: Artificial Neural Network (ANN) and Support Vector Machine (SVM). Both models were evaluated using various feature selection methods to identify the most relevant attributes contributing to accurate classification.

We applied multiple feature selection techniques, including filter, wrapper and embedded methods to reduce the dimensionality of the dataset and enhance model performance. Among these, the wrapper method proved to be the most effective, as it iteratively selected the best subset of features by evaluating model performance during the training phase.

The models were evaluated using standard performance metrics such as accuracy, detection rate, precision, recall and F1- score. The comparative analysis clearly indicated that the ANN model was the most effective in detecting network intrusions.

ACKNOWLEDGEMENT

We extend our sincere gratitude to Mr. M. Rajkumar for his invaluable guidance and support throughout this research. His guidance has been instrumental in shaping the development of this project.

We also thank the faculty members of the Department of Computer Science and Engineering, J.B. institute of Engineering and Technology, for their encouragement and valuable suggestions, which have greatly contributed to the successful completion of this work.

Finally, we acknowledge the contributions of our peers, open-source communities and the creators of essential datasets and tools that made this research possible.

CONCLUSION

In this project, we explored various machine learning models by applying different algorithms and feature selection techniques to identify the most effective model for network intrusion detection. After extensive analysis, we

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

observed that the model built using Artificial Neural Networks (ANN) combined with a wrapper-based feature selection method outperformed all other models. The ANN model achieved an impressive detection rate of 95%, demonstrating its ability to classify network traffic accurately. On the other hand, the Support Vector Machine (SVM) model, despite its robustness, achieved only a 50% detection rate, highlighting the superior performance of ANN. Our approach effectively reduces false positives and enhances overall security by accurately distinguishing between normal and malicious network traffic. The application of feature selection played crucial role in eliminating irrelevant features, thereby improving the model's efficiency and accuracy. Through this project, we have introduced an innovative approach to strengthen cybersecurity by leveraging machine learning techniques for intrusion detection.

REFERENCES

[1] H. Song, M.J. Lynch, and J.K. Cochran, "A macro-social exploratory analysis of the rate of interstate cybervictimization", American journal of Criminal Justice, vol.41, no. three, pp.583-601, 2016.
[2] P. Alaei and F. Noorbehbahani, "Incremental anomaly-based intrusion detection system using limited labelled data", in Web Research (ICWR), 2017 3th International Conference on, 2017, pp. 178-184.
[3] Srinivas Mukkamala, Guadalupe Janoski, Andrew Sung Intrusion Detection: Support Vector Machines and Neural Networks.
[4] Wei Li "Using Genetic Algorithm for Network Intrusion Detection".
[5] Cheng, Tay & Huang, 2012 "Online sequential extreme learning Machine" (OSELM).
[6] Liu, Chen, Liao & Zhang, "Intrusion detection techniques".