IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

# NETWORK INTRUSION DETECTION SYSTEM FOR FEATURE SELECTION - BASED MACHINE LEARNING TECHNIQUE USING ANN & SVM

**Dr.A.Mummoorthy[1] ,**
Associate Professor
Department of Information Technology
Malla Reddy College Of Engineering & Technology
Hyderabad,India.

**Emani Mounika[2],**
Final Year Student
Department of Information Technology
Malla Reddy College Of Engineering & Technology
Hyderabad,India.

**Gaddamidi Sri Sai Charan[3],**
Final Year Student
Department of Information Technology
Malla Reddy College Of Engineering & Technology
Hyderabad,India.

**Gaddam Sri Pavani[4] ,**
Final Year Student
Department of Information Technology
Malla Reddy College Of Engineering & Technology
Hyderabad,India.

## Abstract

This study evaluates performance of two supervised machine learning algorithms such as SVM (Support Vector Machine) and ANN (Artificial Neural Networks). Machine learning algorithms will be used to detect whether request data contains normal or attack (anomaly) signatures. Now-a-days all services are available on internet and malicious users can attack client or server machines through this internet and to avoid such attack request IDS (Network Intrusion Detection System) will be used, IDS will monitor request data and then check if its contains normal or attack signatures, if contains attack signatures then request will be dropped. IDS will be trained with all possible attacks signatures with machine learning algorithms and then generate train model, whenever new request signatures arrived then this model applied on new request to determine whether it contains normal or attack signatures. In this paper we are evaluating performance of two machine learning algorithms such as SVM and ANN and through experiment we conclude that ANN outperform existing SVM in terms of accuracy. To avoid all attacks IDS systems has developed which process each incoming request to detect such attacks and if request is coming from genuine users then only it will forward to server for processing, if request contains attack signatures then IDS will drop that request and log such request data into dataset for future detection purpose. To detect such attacks IDS will be prior train with all possible attacks signatures coming from malicious user's request and then generate a training model. Upon receiving new request IDS will apply that request on that train model to predict it class whether request belongs to normal class or attack class. To train such models and prediction various data mining classification or prediction algorithms will be used. In this study we are evaluating performance of SVM and ANN. In this algorithms we have applied Correlation Based and Chi-Square Based feature selection algorithms to reduce dataset size, this feature selection algorithms removed irrelevant data from dataset and then used model with important features, due to this features selection algorithms dataset size will reduce and accuracy of prediction will increase.

## I. INTRODUCTION

With the wide spreading usages of internet and increases in access to online contents, cybercrime is also happening at an increasing rate [1-2]. Intrusion detection is the first step to prevent security attack. Hence the security solutions such as Firewall, Intrusion Detection System (IDS), Unified Threat Modeling (UTM) and Intrusion Prevention System (IPS) are getting much attention in studies. IDS detects attacks from a variety of systems and network sources by collecting information and then analyzes the information for possible security breaches [3]. The network based IDS analyzes the data packets that travel over a network and this analysis are carried out in two ways. Till today anomaly based detection is far behind than the detection that works based on signature and hence anomaly based detection still remains a major area for research [4-5]. The challenges with anomaly based intrusion detection are that it needs to deal with novel attack for which there is no prior knowledge to identify the anomaly. Hence the system somehow needs to have the intelligence to segregate which traffic is harmless and which one is malicious or anomalous and for that machine learning techniques are being explored by the researchers over the last few years [6]. IDS however is not an answer to all security related problems. For example, IDS cannot compensate weak identification and authentication mechanisms or if there is a weakness in the network protocols.

Studying the field of intrusion detection first started in 1980 and the first such model was published in 1987 [7]. For the last few decades, though huge commercial investments and substantial research were done, intrusion detection technology is still immature and hence not effective [7]. While network IDS that works based on signature have seen commercial success and widespread adoption by the technology based organization throughout the globe, anomaly based network IDS have not gained success in the same scale. Due to that reason in the field of IDS, currently anomaly based detection is a major focus area of research and development [8]. And before going to any wide scale deployment of anomaly based intrusion detection system, key issues remain to be solved [8]. But the literature today is limited when it comes to compare on how intrusion detection performs when using supervised machine learning techniques [9]. To protect target systems and networks against malicious activities anomaly-based network IDS is a valuable technology. Despite the variety of anomaly-based network intrusion detection techniques described in the literature in recent years [8], anomaly detection functionalities enabled security tools are just beginning to appear, and some important problems remain to be solved. Several anomaly based techniques have been proposed including Linear Regression, Support Vector Machines (SVM), Genetic Algorithm, Gaussian mixture model, knearest neighbor algorithm, Naive Bayes classifier, Decision Tree [3,5]. Among them the most widely used learning algorithm is SVM as it has already established itself on different types of problem [10]. One major issue on anomaly based detection is though all these proposed techniques can detect novel attacks but they all suffer a high false alarm rate in general. The cause behind is the complexity of generating profiles of practical normal behavior by learning from the training data sets [11]. Today Artificial Neural Network (ANN) are often trained by the back propagation algorithm, which had been around since 1970 as the reverse mode of automatic differentiation [12].

The major challenges in evaluating performance of network IDS is the unavailability of a comprehensive network based data set [13]. Most of the proposed anomaly based techniques found in the literature were evaluated using KDD CUP 99 dataset [14]. In this paper we used SVM and ANN –two machine learning techniques, on NSLKDD [15] which is a popular benchmark dataset for network intrusion.

## II. EARLIER WORK

The existing Network Intrusion Detection System (NIDS) operates on signature-based detection, employing predefined patterns or signatures of known threats to identify malicious activity within network traffic. This system relies heavily on databases of signatures, continually updated to reflect emerging threats. While effective against known attacks, it struggles with detecting novel or sophisticated threats not covered by existing signatures. Moreover, the NIDS may exhibit high false positive rates, triggering alerts for benign network activities that resemble known attack patterns. False negatives, where genuine threats go undetected, are also a concern, potentially leading to security breaches.

Furthermore, the existing NIDS lacks the ability to adapt to evolving attack techniques and may not effectively detect zero-day exploits or polymorphic malware. Its static nature limits its capacity to recognize anomalous behaviors indicative of new or previously unseen threats. Additionally, the system may face challenges in scaling to handle increasing network traffic volumes or in effectively monitoring encrypted traffic, which modern attackers exploit to evade detection.

Integration with other security tools and processes may be limited, hindering the effectiveness of incident response efforts. The NIDS operates in isolation, without seamless coordination with Security Information and Event Management (SIEM) systems, threat intelligence feeds, or incident response workflows. This lack of integration results in disjointed security operations, slowing down detection and response times.

The maintenance of the existing NIDS could also pose challenges. It requires regular updates to its signature databases to remain effective against evolving threats. However, this process may be manual and time-consuming, potentially leading to delays in threat detection. Moreover, managing the system's configuration and ensuring its compatibility with other security solutions may be complex and resource-intensive.

In summary, the existing NIDS, based on signature-based detection, faces several limitations:

- Inability to detect novel or sophisticated threats not covered by existing signatures.
- High false positive rates and false negatives.
- Lack of scalability to handle increasing network traffic volumes.
- Limited integration with other security tools and processes.
- Maintenance challenges, including manual updates and complex configuration management.

## III. PROPOSED METHOD

The proposed Network Intrusion Detection System (NIDS) aims to overcome the limitations of traditional signature-based detection by leveraging advanced techniques such as Artificial Neural Networks (ANN) and Support Vector Machines (SVM). This next-generation NIDS will enhance threat detection capabilities, reduce false positive rates, and improve overall security posture.

Artificial Neural Networks (ANN) and Support Vector Machines (SVM) are machine learning algorithms well-suited for anomaly detection in network traffic. ANN excels at learning complex patterns and behaviors from large datasets, enabling the detection of novel or previously unseen threats. SVM, on the other hand, is proficient in efficient classification, distinguishing

between normal and malicious network traffic based on learned patterns.

The proposed NIDS will utilize ANN and SVM in a complementary manner to achieve robust threat detection and classification. ANN will be trained on historical network traffic data to learn normal patterns and identify deviations indicative of potential intrusions. SVM will then classify network traffic instances based on the features extracted by ANN, providing an additional layer of detection and reducing false positive rates. Integration with other security tools and processes will be a key component of the proposed NIDS. It will seamlessly integrate with Security Information and Event Management (SIEM) systems, threat intelligence feeds, and incident response workflows. This integration will facilitate swift detection and response to security incidents by providing contextual information and enabling automated incident triage and remediation.

The proposed NIDS will be designed for scalability and ease of management. It will feature streamlined configuration processes and automated updates to ensure timely adaptation to evolving threats. Additionally, the system will be capable of handling increasing network traffic volumes and effectively monitoring encrypted traffic, leveraging advanced algorithms optimized for performance and scalability.

The implementation plan for the proposed NIDS includes several key steps:

**Comprehensive Assessment:** Conduct a thorough assessment of existing NIDS capabilities and identify gaps in threat detection and response.

**Requirements Definition:** Define requirements and objectives for the proposed NIDS, considering factors such as detection accuracy, false positive rates, scalability, and integration capabilities.

**Algorithm Selection:** Select appropriate machine learning algorithms, tools, and vendors based on requirements and desired outcomes.

**Architecture Design:** Design the architecture and deployment strategy for the NIDS, including data preprocessing, feature selection, model training, and deployment of sensors.

**Configuration and Deployment:** Configure and deploy the NIDS in phases, starting with critical network segments and gradually expanding coverage.

**Testing and Validation:** Conduct thorough testing and validation to ensure the effectiveness and reliability of the NIDS in detecting and responding to security threats.

**Training and Education:** Train security personnel on the operation and management of the NIDS, as well as the interpretation of results and integration with other security tools and processes.

**Monitoring and Evaluation**: Monitor and evaluate the performance of the NIDS over time, making adjustments as necessary to optimize detection capabilities and response times.

## IV. METHODOLOGY

**1. Define Objectives and Requirements:**

Begin by clearly defining the objectives of your NIDS deployment. Identify the assets you need to protect, potential threats, and compliance requirements. Determine the scope and scale of your NIDS deployment and establish performance requirements.

**2. Network Architecture Analysis:**

Understand your network architecture thoroughly. Identify critical network segments, entry points, and potential vulnerabilities. This analysis will help in determining where to deploy sensors effectively.

**3.Sensor Placement:**

Based on the network architecture analysis, decide where to place NIDS sensors strategically. This includes placing sensors at network boundaries, critical chokepoints, and within internal segments where sensitive data resides. Ensure comprehensive coverage while considering performance and scalability.

**4. Select NIDS Technology:**

Choose an appropriate NIDS technology based on your requirements, such as signature-based, anomaly-based, or behavior-based detection. Consider factors like scalability, performance, ease of management, and integration capabilities with existing security infrastructure.

**5. Configuration and Tuning:**

Configure NIDS sensors according to best practices and customize detection rules to align with your organization's security policies and specific threat landscape. Fine-tune detection thresholds to minimize false positives and false negatives.

**6. Integration with Security Operations:**

Integrate NIDS with other security tools and processes such as SIEM (Security Information and Event Management) systems, incident response procedures, and threat intelligence feeds. Ensure seamless coordination for effective incident detection, analysis, and response.

**7. Continuous Monitoring and Maintenance:**

Regularly monitor NIDS alerts and performance metrics to detect and respond to potential security incidents promptly. Keep NIDS signatures and detection capabilities up-to-date to defend against emerging threats. Perform routine maintenance tasks like software updates, sensor health checks, and periodic configuration reviews.

**8.Incident Response and Remediation:**

Establish clear incident response procedures for handling NIDS alerts and security incidents. Define roles and responsibilities, escalation paths, and communication protocols. Develop playbooks for responding to common security incidents detected by NIDS and conduct regular tabletop exercises to validate the effectiveness of incident response processes.

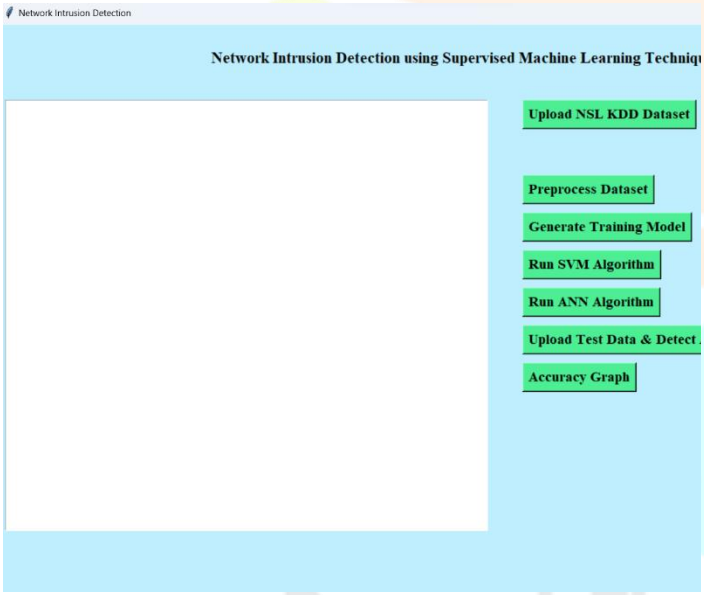
**9.Performance Evaluation and Optimization:**

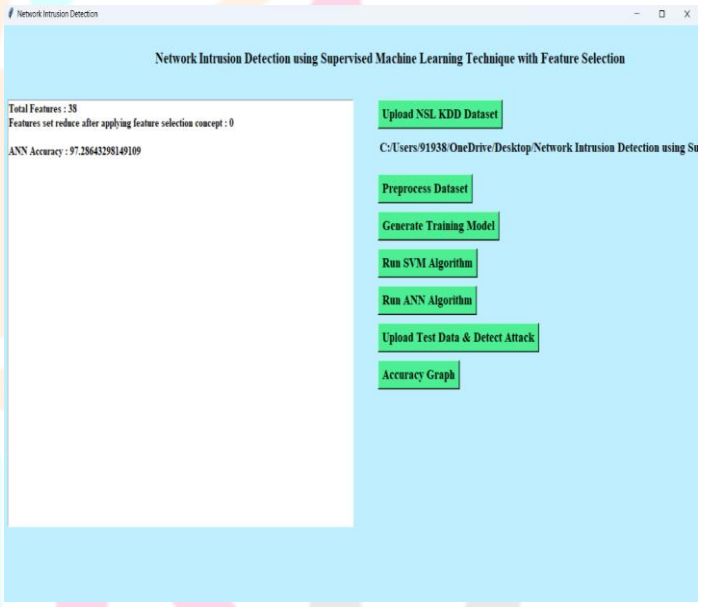
Periodically evaluate the performance and effectiveness of your NIDS deployment against predefined objectives and key performance indicators (KPIs). Identify areas for improvement and optimization, such as enhancing detection capabilities, optimizing sensor placement, or upgrading hardware/software components to scale with evolving network requirements.

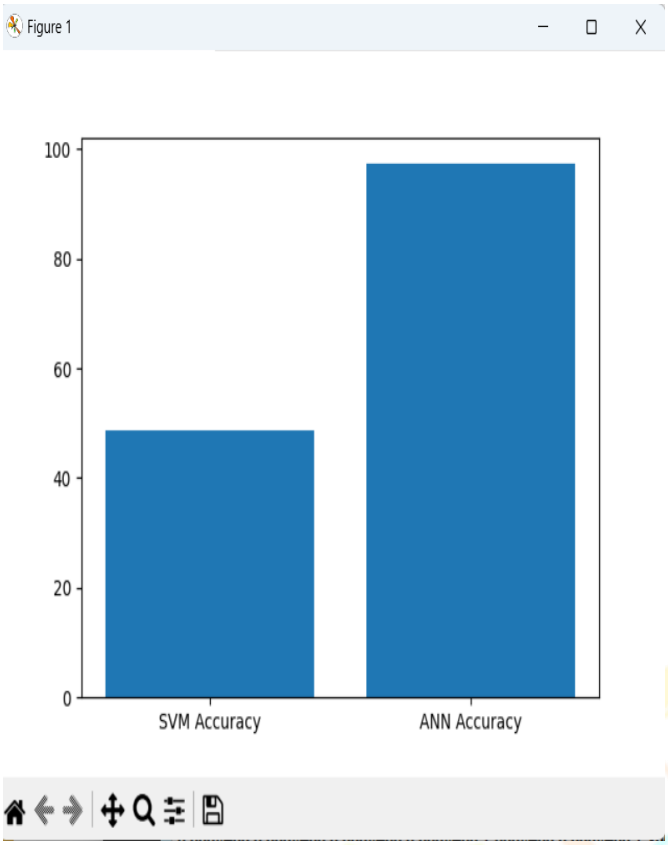
**10.Compliance and Reporting:**

Ensure compliance with relevant regulations and industry standards governing network security and intrusion detection. Generate regular reports summarizing NIDS activities, including detected incidents, response actions taken, and recommendations for enhancing security posture.


## RESULT AND ANALYSIS






## VI CONCLUSION

This study worked on issue related to Naïve Bayes machine learning classifier as it assume strong feature independence between attributes so proposed new algorithm which approximates the interactions between attributes by using conditional probabilities. The performance comparison amongst different classifiers with proposed classifier is made in order to understand their effectiveness in terms of various performance measures. From results, it is clear that every attributes in data set is not of equal importance, as we can ignore some attributes over others which does not involve much in intrusion detection. So this study has applied the feature selection techniques and found better results than before. Experimental result illustrates feature subset identified by Gain ratio + Ranker has improved our proposed Naïve Bayes classification. In future we will try to implement feature selection using soft computing techniques to identify intrusion in adaptive heterogeneous environment.

## VII. REFERENCES

1)Chih-Fong Tsai a, Yu-Feng Hsu b, Chia-Ying Lin c, Wei-Yang Lin d "Intrusion detection by machine learning A review" Expert Systems with Applications Elsevier 2009

2)Tanya Garg and Surinder Singh Khurana IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014), May 09-11, 2014, Jaipur, India

3)Jian Pei Shambhu J. Upadhyaya Faisal Farooq Venugopal Govindaraju. Proceedings of the 20thInternational Conference on Data Engineering published In IEEE 2004.

4)Siva S. Sivatha Sindhu, Geetha , A. Kannan ” Decision tree based light weight intrusion detection using a wrapper approach “.Expert Systems with Applications 39 (2012) 129–141 published in Elsevier

5)Muamer N. Mohammada, Norrozila Sulaimana, Osama Abdulkarim Muhsin “A Novel Intrusion Detection System by using Intelligent Data Mining in Weka Environment”. Procedia Computer Science 3 (2011) 1237–1242

6)F. Maggi, M. Matteucci and S. Zanero, “Reducing false positives in anomaly detectors through fuzzy alert aggregation”. Information Fusion, 10, 300–311. 2009

7)Dr. Saurabh Mukherjee, Neelam Sharma, “Intrusion Detection using Naive Bayes Classifier with Feature Reduction” Published by Elsevier 2012.

8)ENGEN, “Machine learning for network based intrusion detection,” Doctoral dissertation, Bournemouth University, 2010.

9)Paxson, Vern, Bro, “A System for Detecting Network

Intruders in Real-Time,” Proceedings of The 7th USENIX Security Symposium, San Antonio TX, 1998.

10)PAT LANGLEY, STEPHANIE SAGE,” Induction of Selective Bayesian Classifiers” Institute for the Study of Learning and Expertise 2451 High Street, Palo Alto, CA 94301