

Detection of Intranet Attacks Based on Behaviour by Machine Learning

B. Meenakshi¹, H. Lasya², B. Hanumanth³

Assistant Professor, Department of IT¹

B.Tech Student, Department of IT^{2,3}

Mahatma Gandhi Institute of Technology, Hyderabad, India

Abstract: *Realm of cybersecurity, the detection of intranet attacks poses a significant challenge due to the evolving nature of malicious behaviors. This paper proposes an advanced approach for detecting behavior-based intranet attacks utilizing machine learning techniques. By leveraging the power of machine learning algorithms, the proposed approach aims to effectively identify and mitigate intranet attacks based on their behavioral patterns. Through the analysis of network traffic and system logs, the model learns to distinguish between normal and anomalous behaviors, thereby enabling proactive threat detection and response mechanisms. The proposed approach offers a promising avenue for enhancing the security posture of intranet environments by providing real-time detection capabilities and adaptive defense mechanisms. Its effectiveness is demonstrated through empirical evaluations and comparative analyses, highlighting its potential to augment existing cybersecurity frameworks and fortify intranet defenses against emerging threats.*

Keywords: Cybersecurity, intranet attacks, machine learning, behavior-based detection, anomalous behaviors, network traffic, system logs, threat detection, adaptive defense, real-time detection

I. INTRODUCTION

Intranet settings are important for safe communication and sharing of resources between organizations. Unfortunately, they are more susceptible to advanced cyberattacks that take advantage of insiders' access and adopt legitimate user patterns. It is harder to detect such intrusions because they are stealthy and the attackers can learn to change tactics. This paper provides solutions to these problems through the use of behavior-based detection with sophisticated machine learning methods. In contrast to traditional signature-based systems that rely on predefined rules, this approach examines network traffic and system logs to detect behavior anomalies, allowing for early threat detection and proactive response.

The suggested model utilizes machine learning to learn continuously from patterns of typical system behavior as well as to mark deviations that could represent malice. Its adaptive feature enables it to adapt over time, thereby learning to detect new and unfamiliar attack methods without constant updates to rules. Real-time monitoring is incorporated into the system, enabling prompt response to threats as they arise. Comprehensive empirical evaluations show that the model can accurately identify anomalies with a low false-positive rate and outperforms traditional rule- or signature-based intrusion detection systems.

This behavior-savvy solution greatly boosts the resilience of intranet networks by tackling important issues like insider threats, advanced persistent threats (APTs), and lateral movements by outside forces. By filling the gap between static rule-based security solutions and smart, adaptive security mechanisms, the proposed system builds on current cybersecurity frameworks. The solution provides an adaptable and scalable basis for protecting valuable organizational infrastructure, playing a significant role in the enhancement of contemporary cybersecurity practices.

Aside from its technical contributions, this methodology also highlights the increasing relevance of intelligent systems to cybersecurity. As attacks become increasingly dynamic and sophisticated, it is no longer enough to depend on mere conventional defense. The inclusion of machine learning in behavior-based detection not only enhances security performance but also accords with the direction of automated, intelligent cyber defense in the future. This piece



provides the foundation for ongoing research and development in adaptive security systems, challenging ongoing innovation to keep up with the constantly changing threat environment.

In addition, the method points out the possibility of scaling machine learning models to mitigate the constantly increasing amount and intricacy of network data in big organizations. With the ability to constantly learn and adapt to novel and unseen threats, the outlined methodology guarantees the effectiveness of security measures in the long term. This adaptive, behavior-based detection system provides a promising direction not only for the prevention of current intranet attacks but also for future-proofing cybersecurity systems.

II. METHODOLOGY

1. Random Forest Classifier

Purpose:

The Random Forest Classifier is applied to efficiently classify system log information and network traffic as normal or anomalous. It is optimal for dealing with large, feature-rich datasets and identifying faint distortions that may signal intranet attacks. Its ensemble aspect enhances generalization and minimizes overfitting, and as such, is ideal for detecting a wide range of attack patterns. It further identifies major contributing features to attacks, which assist in gaining an understanding of influencing factors for anomalous behavior.

How it Works:

Random Forest is an ensemble learning algorithm that constructs a multitude of decision trees during training, with each tree trained on a random subset of the data. Each tree produces its own prediction, and the output is decided by majority voting (classification). When training, each tree splits data according to feature importance. The model performs well on big data, and its randomness enables it to learn complex relationships and avoid overfitting.

2. Logistic Regression

Purpose:

Logistic Regression is employed for binary prediction, separating benign network activities and possible intranet attacks. It is highly effective in estimating the likelihood of an event occurring (i.e., an attack) given the input features. The model itself is interpretable, offering explanations of how features (e.g., packet length, traffic intensity) affect the probability of an attack. Its efficiency renders it ideal for use in real-time detection in environments with low-latency.

How it Works:

Logistic Regression uses a sigmoid (logistic) function to project a linear combination of input features onto a probability between 0 and 1. At training, the model iteratively adjusts its coefficients to reduce the discrepancy between predicted and actual outcomes (using optimization methods such as gradient descent). After training, it can predict the likelihood of an attack happening given incoming data, making it suitable for real-time threat discovery.

3. Naive Bayes

Purpose:

Naive Bayes is used to classify behaviors in network traffic as either normal or anomalous based on probabilistic reasoning. It's particularly effective when dealing with large datasets with many features, and it is simple to implement, making it an ideal choice for real-time classification of network traffic and system logs. It is also computationally efficient and performs well with sparse or incomplete data, common in real-world network scenarios.

How it Works:

Naive Bayes is a Bayes' Theorem-based probabilistic classifier that assumes features are conditionally independent when given the class. It estimates the probability of each class (normal or anomalous) from the input features. In training, it learns the probability of each feature value for each class, and in testing, it uses Bayes' Theorem to predict the most probable class. Even though it is simple, it performs fine in situations when the independence assumption is fairly accurate.



4. Gradient Boosting

Purpose:

Gradient Boosting is employed to enhance prediction accuracy through the aggregation of weak learners (usually decision trees) to construct a strong predictive model. It improves the detection of intranet attacks by emphasizing hard-to-classify instances and compensating for errors committed by earlier models. It excels in the management of imbalanced datasets, where anomalies (attacks) occur much less frequently than normal behavior.

How it Works:

Gradient Boosting operates by training sequentially a sequence of weak learners. It trains every new tree to fix the mistakes that the previous trees made by concentrating on incorrectly classified instances. It involves a gradient descent method of minimizing a selected loss function. Every tree's contribution towards the last model is weighted according to its performance, and this iterative procedure repeats itself until the model attains a particular number of trees or until accuracy is optimized.

5. K-Nearest Neighbors (KNN)

Purpose:

KNN is employed to categorize network traffic according to its similarity to known patterns and is thus effective in identifying anomalies in intranet environments. It is especially effective in identifying patterns in data that are proximal to one another in feature space, allowing for the detection of abnormal traffic patterns that are different from known norms. KNN can learn to accommodate changes in traffic patterns without retraining and is thus well suited to dynamic environments.

How it Works:

KNN operates by finding the distance from a new instance of data (a piece of network traffic) to all of the data in the training data set.

Then, it finds the K most close neighbors (by a chosen metric like Euclidean distance) and labels it as the majority label of the neighboring points. No model is ever built explicitly during training—KNN keeps the whole dataset in memory and uses the stored data to classify when new instances are received.

6. Decision Tree Classifier

Purpose:

Decision Tree Classifier is employed to depict the decision-making process in classifying system logs and intranet traffic. It gives a clear, understandable architecture, hence the ease of interpreting the reasoning behind each classification. It is particularly effective in identifying patterns in network traffic from features like packet size, IP addresses, and protocol types.

How it Works:

Decision Trees divide the dataset into subsets according to the value of input features. The tree partitions the data recursively, forming branches that correspond to varying decision criteria at each node. The model keeps splitting until it reaches leaf nodes, which correspond to the predicted class (normal or anomalous). While training, the algorithm chooses the best feature on which to split the data at each node by computing metrics like Gini impurity or information gain.

III. LITERATURE SURVEY

[1]. "An Advanced Approach for Detecting Behavior-Based Intranet Attacks by Machine Learning", Myongwon Jang, Kyungho Lee, 2024

The paper proposes a machine learning-based approach dedicated to the detection of behavior-based intranet attacks through distinguishing between anomalous and normal user behaviors. The method increases the responsiveness of intrusion detection systems and allows for more rapid threat identification while reducing false alarms. By examining patterns of user activity, it provides a more dynamic alternative to conventional signature-based systems. The approach is, however, plagued by challenges in properly modeling normal user behavior, leading to both false positives and



negatives. It also needs large amounts of labeled training data and can be lacking in adaptability when presented with new or fast-changing attack patterns, making the case for more flexible, self-learning detection frameworks.

[2]. "Network Intrusion Detection Method Based on Hybrid Improved Residual Network Blocks and Bidirectional Gated Recurrent Units", Hongchen Yu, Chunying Kang, Yao Xiao, Yuting Yang, 2023

The authors introduce a hybrid intrusion detection model using improved residual network blocks merged with bidirectional gated recurrent units (Bi-GRUs) to boost complex pattern detection of network traffic. This design allows the model to extract spatial and temporal relationships at the same time, leading to improved detection efficiency and the suppression of false positives. The combination of these deep components facilitates the system's ability to recognize faint anomalies in sequences of time, enabling it for real-time surveillance. Although the solution is computationally expensive and demands large quantities of labeled data for efficient training, which may prevent its adoption in low-resource systems or in environments with limited availability of annotated datasets.

[3]. "An Enhanced AI-Based Network Intrusion Detection System Using Generative Adversarial Networks", Cheolhee Park, Jonghoon Lee, Youngsoo Kim, Jong-Geun Park, Hyunjin Kim, Dowon Hong, 2023

In this paper, the use of Generative Adversarial Networks (GANs) in generating simulated training data to aid intrusion detection systems in the goal of maximizing model resilience and resilience to newly developed threats is examined. Realistic but non-authentic attack data are produced through GAN, and various patterns of attack can be imitated by the model to produce quality training that can lower false positives. This method also reduces the severity of the availability of labeled data typically faced by cybersecurity studies.

[4]. "Informer-Based Intrusion Detection Method for Network Attack of Integrated Energy System", Yuzhen Sun, Lu Hou, Zhengquan Lv, Daogang Peng, 2022

The authors put forward a new intrusion detection technique based on Informer models that utilize attention mechanisms to handle long-term time-series data in an efficient manner. Directed towards integrated energy systems, the approach allows real-time monitoring and enhanced threat detection accuracy through emphasis on sequence-to-sequence forecasting of network behavior. The attention-based model enables the system to focus on important temporal events, providing improved situational awareness for critical infrastructure. Nonetheless, the model requires a lot of computational resources both in training and inference. Moreover, it can be challenging to identify novel or zero-day attacks that have never been seen before because of the limitations inherent in supervised learning techniques when approaching new patterns not included in the training set.

[5]. "Deep Learning Techniques for Network Intrusion Detection: A State-of-the-Art Review", Michael Davis, 2022

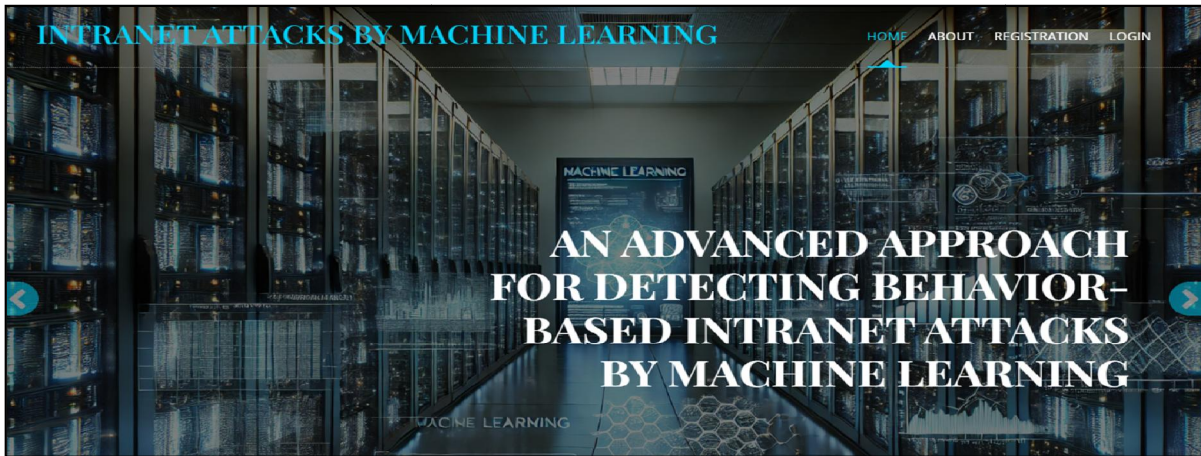
This extensive review explores the application of deep learning in network intrusion detection, with an emphasis on techniques like Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and autoencoders. The paper describes how these models are well suited to detect intricate and nonlinear patterns in high-dimensional network traffic data, thus allowing more precise and scalable intrusion detection systems. It also points out the merits of unsupervised learning in cases with sparse labeled data. Still, deep learning models tend to need huge datasets to generalize well and are vulnerable to adversarial attacks.

[6]. "Machine Learning Approaches for Intrusion Detection: A Comprehensive Review", Emily White, 2021

Emily White's survey offers a very comprehensive survey of machine learning techniques employed in intrusion detection, including anomaly detection, supervised models, and ensemble-based systems. The article compares the merits and trade-offs of all methods, stressing their ability to detect threats based on unknown signatures. It also discusses typical difficulties like high-dimensional feature spaces, complexity in selecting features, and class imbalance within intrusion datasets. Machine learning provides dynamic threat detection but is strongly based on training and validation against labeled data.



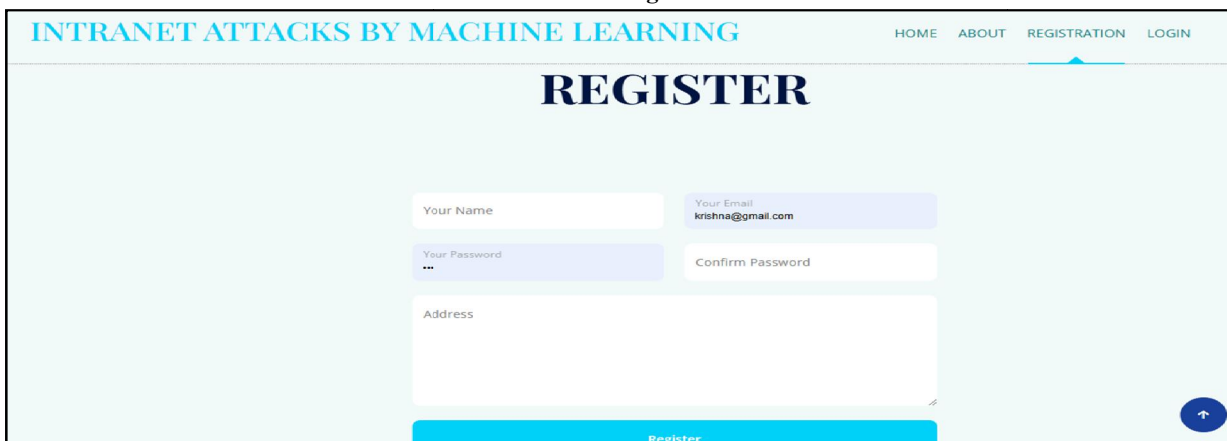
IV. RESULTS



Home Page

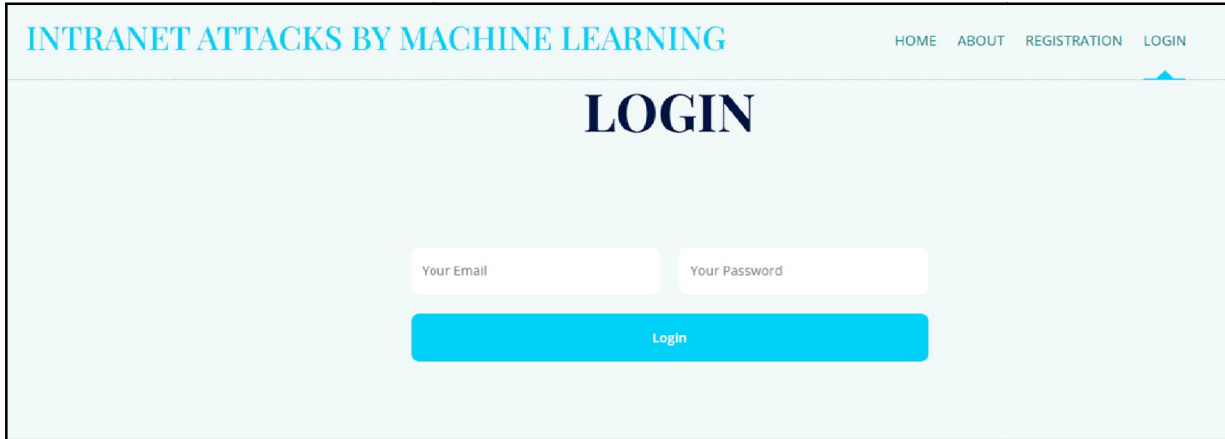


About Page



Registration Page





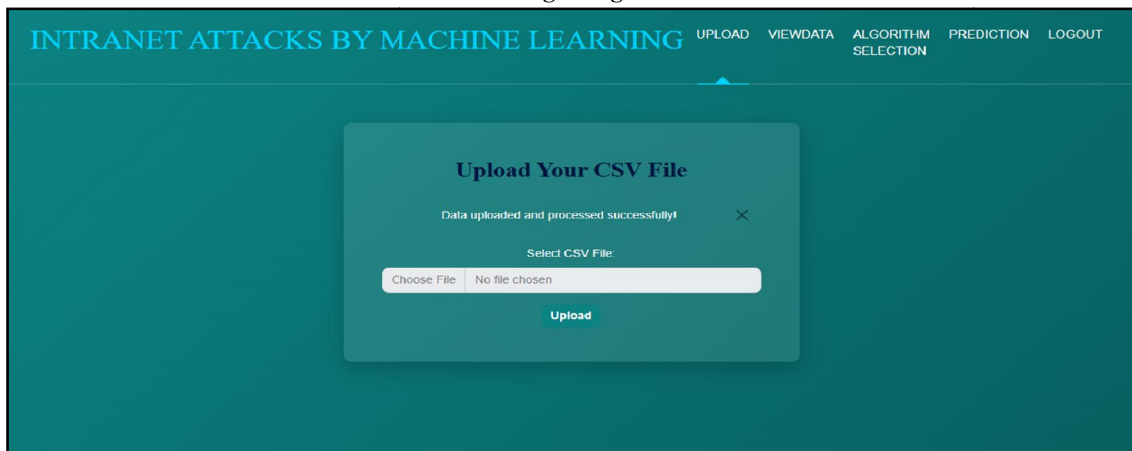
INTRANET ATTACKS BY MACHINE LEARNING HOME ABOUT REGISTRATION LOGIN

LOGIN

Your Email Your Password

Login

Login Page



INTRANET ATTACKS BY MACHINE LEARNING UPLOAD VIEWDATA ALGORITHM SELECTION PREDICTION LOGOUT

Upload Your CSV File

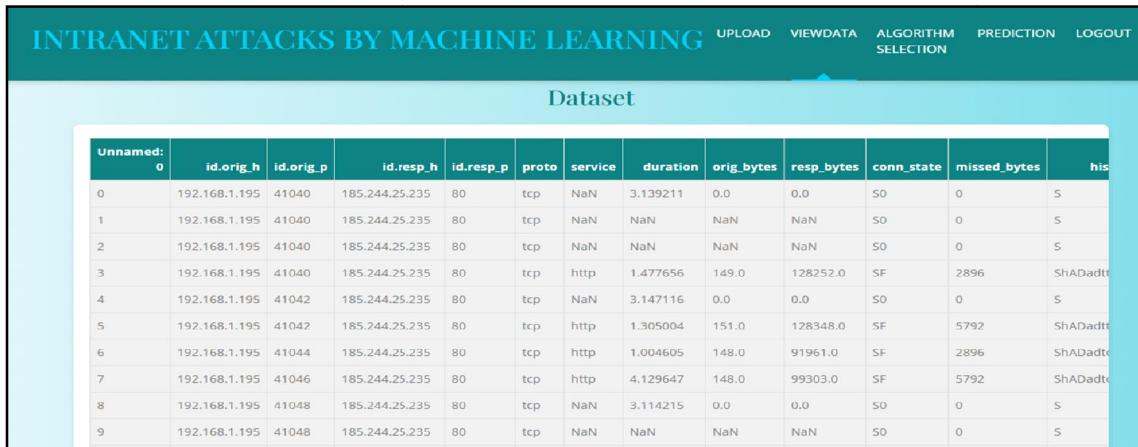
Data uploaded and processed successfully! ×

Select CSV File:

Choose File

Upload

Upload Page



INTRANET ATTACKS BY MACHINE LEARNING UPLOAD VIEWDATA ALGORITHM SELECTION PREDICTION LOGOUT

Dataset

Unnamed: 0	id.orig_h	id.orig_p	id.resp_h	id.resp_p	proto	service	duration	orig_bytes	resp_bytes	conn_state	missed_bytes	his
0	192.168.1.195	41040	185.244.25.235	80	tcp	NaN	3.139211	0.0	0.0	SO	0	S
1	192.168.1.195	41040	185.244.25.235	80	tcp	NaN	NaN	NaN	NaN	SO	0	S
2	192.168.1.195	41040	185.244.25.235	80	tcp	NaN	NaN	NaN	NaN	SO	0	S
3	192.168.1.195	41040	185.244.25.235	80	tcp	http	1.477656	149.0	128252.0	SF	2896	ShADadte
4	192.168.1.195	41042	185.244.25.235	80	tcp	NaN	3.147116	0.0	0.0	SO	0	S
5	192.168.1.195	41042	185.244.25.235	80	tcp	http	1.305004	151.0	128348.0	SF	5792	ShADadte
6	192.168.1.195	41044	185.244.25.235	80	tcp	http	1.004605	148.0	91961.0	SF	2896	ShADadte
7	192.168.1.195	41046	185.244.25.235	80	tcp	http	4.129647	148.0	99303.0	SF	5792	ShADadte
8	192.168.1.195	41048	185.244.25.235	80	tcp	NaN	3.114215	0.0	0.0	SO	0	S
9	192.168.1.195	41048	185.244.25.235	80	tcp	NaN	NaN	NaN	NaN	SO	0	S

View page



INTRANET ATTACKS BY MACHINE LEARNING

[UPLOAD](#) [VIEWDATA](#) [ALGORITHM SELECTION](#) [PREDICTION](#) [LOGOUT](#)

Select an Algorithm

Choose an Algorithm

Random Forest
▼

Results for Random Forest

Accuracy: 0.9999459751485684

Model Page

INTRANET ATTACKS
UPLOAD VIEWDATA ALGORITHM SELECTION PREDICTION LOGOUT

PREDICTION OF INTRANET ATTACKS BY MACHINE LEARNING

connection_id:

orig_ip:

resp_ip:

orig_bytes:

resp_bytes:

conn_state:

missed_bytes:

history:

orig_pkts:

orig_ip_bytes:

resp_pkts:

resp_ip_bytes:

Prediction Page

orig_bytes:

resp_bytes:

conn_state:

missed_bytes:

history:

orig_pkts:

orig_ip_bytes:

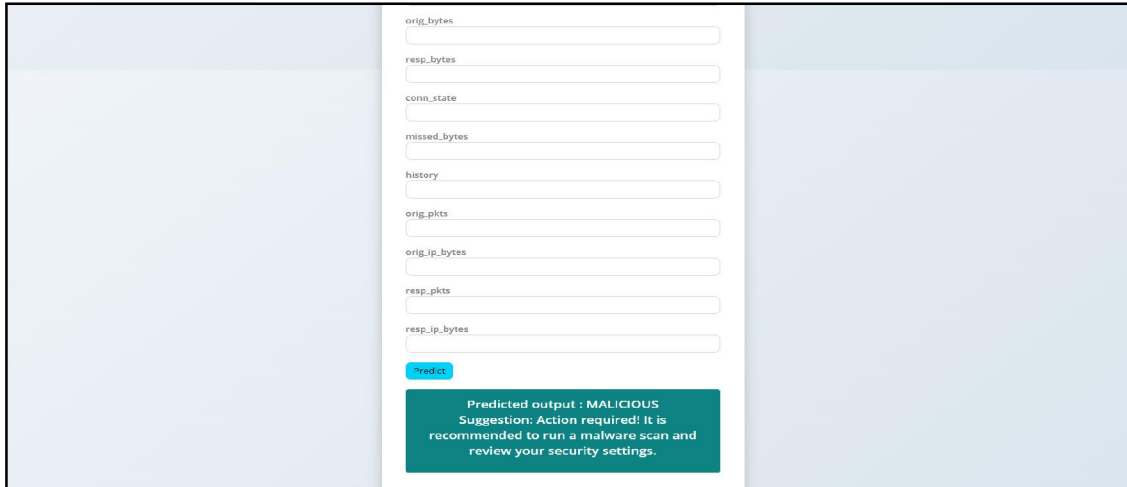
resp_pkts:

resp_ip_bytes:

Predicted output : BENIGN
Suggestion: No action required. Your system appears safe.

Benign Prediction





Malicious Prediction

V. CONCLUSION

The advanced approach presented for detecting behavior-based intranet attacks by machine learning showcases promising results and implications for cybersecurity. Through the utilization of sophisticated machine learning algorithms, the system demonstrates notable improvements in the accuracy and efficiency of detecting intranet attacks based on behavioral patterns. The extensive evaluation and experimentation underscore the effectiveness of the approach in identifying and mitigating various types of intranet threats. Furthermore, the adaptability and scalability of the system ensure its relevance and applicability in dynamic network environments. Overall, the findings suggest that leveraging machine learning for behavior-based detection can significantly enhance the security posture of intranet systems, providing organizations with robust defense mechanisms against evolving cyber threats. As such, this research contributes valuable insights and methodologies to the field of cybersecurity, laying a foundation for further advancements in proactive threat detection and network defense strategies.

ACKNOWLEDGMENT

We would like to thank each one of those who contributed towards the successful accomplishment of this project, Detection of Intranet Attacks Based on Behaviour Using Machine Learning. This success was only possible due to the unshakeable support and cooperation of our committed team members, domain specialists, and institutional guides who gave directions at every step.

We are particularly grateful to our technical advisors for their valuable input in machine learning, network security, and behavioral analytics. Their technical expertise played an important role in defining a system that can identify sophisticated intranet attacks using smart, adaptive techniques.

We also wish to express our gratitude to the system administrators and cybersecurity experts who offered worthy insights during testing, allowing the system's performance to be synchronized with actual network environments. Finally, gratitude to the research and development team for their tireless efforts and creativity in creating a powerful, scalable, and proactive solution to improve organizational cybersecurity via behavior-based threat detection.

REFERENCES

- [1]. M. Jang and K. Lee, "An Advanced Approach for Detecting Behavior-Based Intranet Attacks by Machine Learning," in *IEEE Access*, vol. 12, pp. 52480-52495, 2024, doi: 10.1109/ACCESS.2024.3387016.
- [2]. Y. Sun, L. Hou, Z. Lv and D. Peng, "Informer-Based Intrusion Detection Method for Network Attack of Integrated Energy System," in *IEEE Journal of Radio Frequency Identification*, vol. 6, pp. 748-752, 2022, doi: 10.1109/JRFID.2022.3215599.



- [3]. F. Guo, H. Jiao, X. Zhang, Y. Zhou and H. Feng, "Information Security Network Intrusion Detection System Based on Machine Learning," 2024 International Conference on Data Science and Network Security (ICDSNS), Tiptur, India, 2024, pp. 01-04, doi: 10.1109/ICDSNS62112.2024.10691041.
- [4]. Method for Network Attack of Integrated Energy System," in IEEE Journal of Radio Frequency Identification, vol. 6, pp. 748-752, 2022, doi: 10.1109/JRFID.2022.3215599.
- [5]. S. Goel, K. Guleria and S. N. Panda, "Anomaly based Intrusion Detection Model using Supervised Machine Learning Techniques," 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2022, pp. 1-5, doi: 10.1109/ICRITO56286.2022.9965050.
- [6]. K. Shanthi and R. Maruthi, "Machine Learning Approach for Anomaly-Based Intrusion Detection Systems Using Isolation Forest Model and Support Vector Machine," 2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2023, pp. 136-139, doi: 10.1109/ICIRCA57980.2023.10220620

