International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 06 | June - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

# Detection of Cyber Attack and Network Attack Using Machine Learning

Mr. Ajit Wagh, Mr. Sanket Wandhekar, Mr. Nilesh Wable, Mr. Ravindra Pawar, Prof. M. S. Dighe

**Email:** ajitwagh1787@gmail.com

## ABSTRACT

This study delves into the escalating challenge of cybercrime by highlighting the limitations of current intrusion detection systems (IDS) and the emergent importance of machine learning in bolstering cyber security. Focused on enhancing detection capabilities, the research proposes a novel approach utilizing deep neural networks within a leader-follower framework to identify attacks early in their initiation. Moreover, the study explores resilient control algorithms and reputation mechanisms to isolate and neutralize malicious agents within complex cyber-physical systems (CPS). Experimentation showcases the superior efficacy of deep learning algorithms in detecting threats compared to conventional methods, paving the way for a more proactive, cost-effective, and efficient cyber security landscape.

**Keywords:** Machine Learning, Privilege Escalation, Insider Attack, Classification, TF-IDF, Email etc.

## Introduction

A Develop an innovative cyber attack detection system utilizing machine learning algorithms. The project involves collecting and preprocessing diverse data sources such as network traffic logs and system behaviors. Leveraging supervised and unsupervised learning models, the system will classify normal and malicious activities in real-time. Implementation includes deploying the system for continuous monitoring, triggering alerts upon detecting anomalies, and integrating it with existing security infrastructure. The goal is to create a robust, adaptive system that enhances early threat detection, reduces false positives, and fortifies overall cybersecurity posture against evolving cyber-attacks.

## METHODOLOGIES

Data Preprocessing: Cleaning and preparing raw data to make it suitable for analysis. This includes removing duplicates, handling missing values, and converting data into a usable format for machine learning models.

Feature Engineering: Creating new features or modifying existing ones to improve the model's performance. This might involve extracting relevant information from raw data, such as the number of login attempts or unusual patterns in network traffic.

Normalization and Scaling: Adjusting the range of data values to ensure consistency. Normalization scales the data to a range of 0 to 1, while scaling adjusts data to have a mean of 0 and a standard deviation of 1. This helps algorithms process data more effectively.

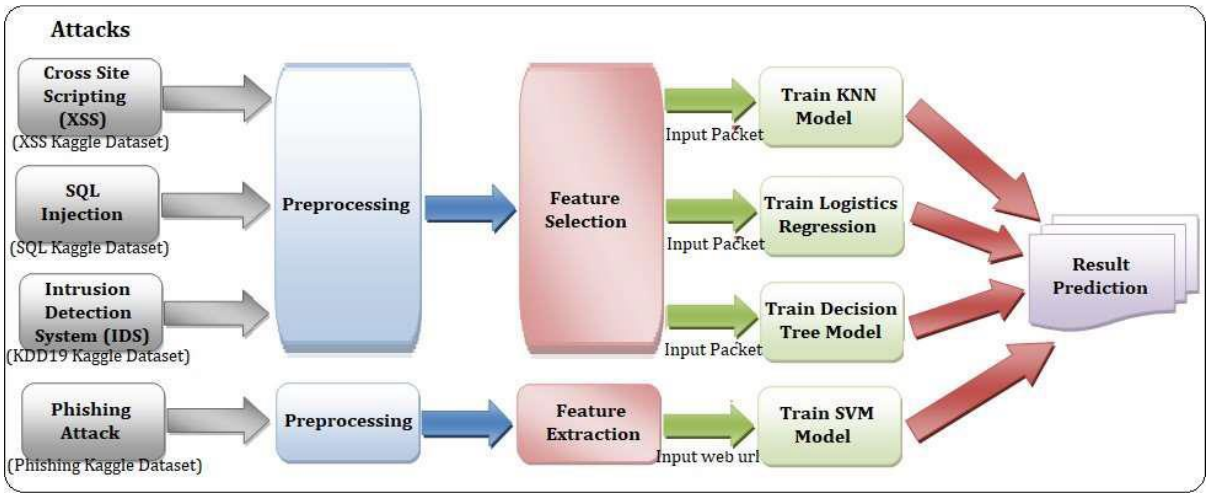
Algorithm Selection: Choosing the most suitable machine learning algorithms for detecting cyber and network attacks. Common algorithms include decision trees, random forests, and neural networks, depending on the nature and complexity of the data.

Supervised Learning: Training the model using labelled data where the outcome (e.g., attack or no attack) is known. The model learns to associate input features with the correct output to predict future attacks.

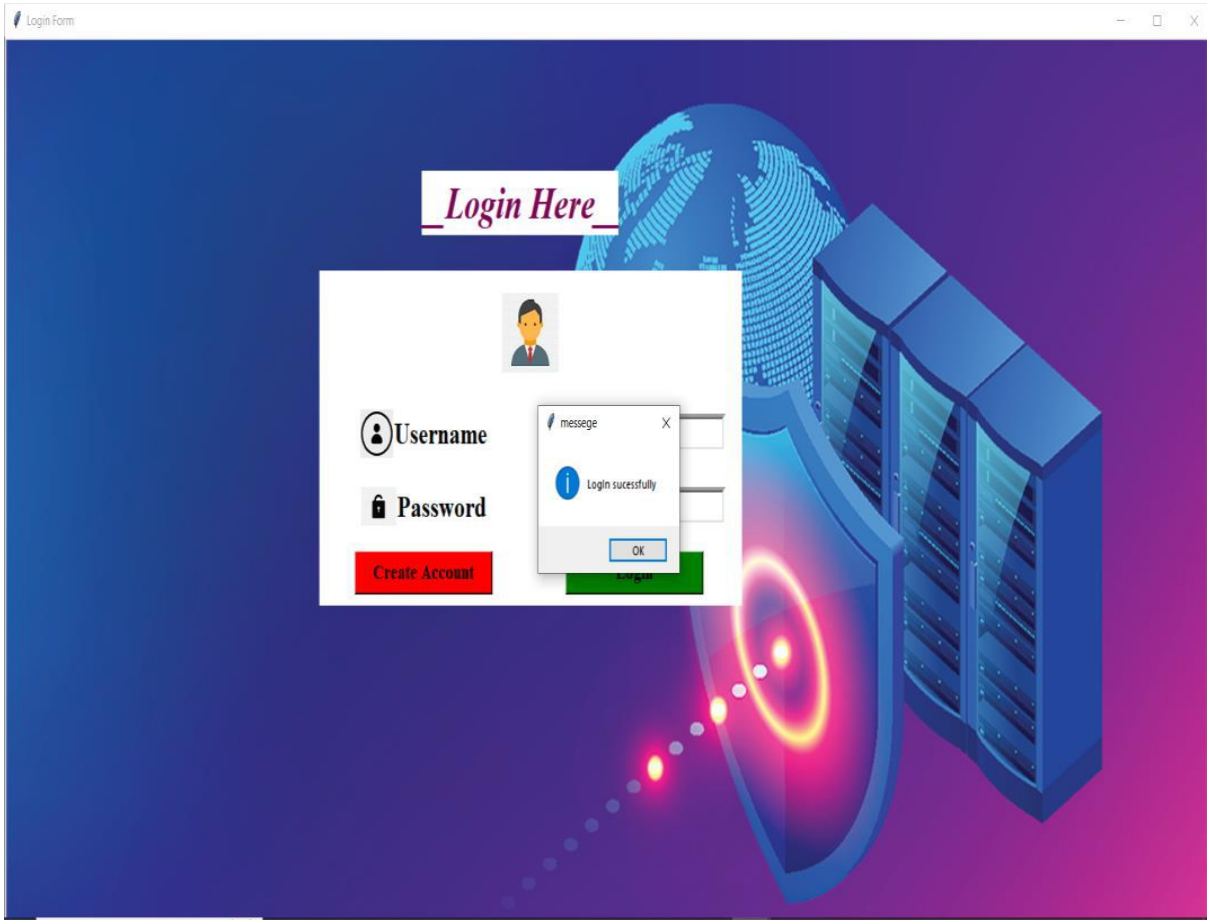
Unsupervised Learning: Analysing data without labelled outcomes to identify patterns or anomalies that may indicate an attack. Techniques like clustering can group similar data points together, revealing unusual behaviour that deviates from the norm.

**System Architecture**

Machine learning techniques address major cybersecurity challenges such as intrusion detection, malware classification, spam detection, and phishing detection. While ML cannot fully automate cybersecurity, it enhances threat identification, easing the burden on security analysts. Efficient adaptive methods like ML result in higher detection rates, lower false alarms, and reasonable computation costs. Detecting attacks is fundamentally harder, making it challenging for intrusion detection to effectively employ ML. ML algorithms can train systems to detect cyber-attacks, triggering email alerts for security users. For instance, classification algorithms like Support Vector Machines (SVM) can categorize attacks, such as DoS/DDoS. Since absolute prevention is impossible, early detection helps mitigate the risk of damage. Organizations can use or develop solutions for early-stage attack detection, ideally with minimal human intervention.


Attacks
Cross Site Scripting (XSS)
(XSS Kaggle Dataset)
SQL Injection
(SQL Kaggle Dataset)
Intrusion Detection System (IDS)
(KDD19 Kaggle Dataset)
Phishing Attack
(Phishing Kaggle Dataset)
Preprocessing
Preprocessing
Feature Selection
Feature Extraction
Input Packet
Input Packet
Input Packet
Input web url
Train KNN Model
Train Logistics Regression
Train Decision Tree Model
Train SVM Model
Result Prediction

## Result


Login Form
_Login Here_
Username
Password
Create Account
messege
Login sucessfully
OK


GUI
Intrusion Detection
Data_Preprocessing
Data Loaded=>Splitted into 80% for Training & 20% for Testing
detection
Exit


Intrusion Detection
Flow_Duration 15842
Total_Fwd_Packets 24
Total_Backward_Packets 22
Total_Length_of_Fwd_Packets 703
Total_Length_of_Bwd_Packets 564
Fwd_Packet_Length_Max 453
Fwd_Packet_Length_Min 0
Fwd_Packet_Length_Mean 667
Fwd_Packet_Length_Std 529
Bwd_Packet_Length_Max 1418
Bwd_Packet_Length_Min 0
Bwd_Packet_Length_Mean 455
Bwd_Packet_Length_Std 105
Flow_Bytes 305
Flow_Packets 305
BENIGN
Submit

## CONCLUSION

The industrial Intrusion network based network is rapidly growing in the coming future. The detection of software piracy and malware Intrusion are the main challenges in the field of cybersecurity using Intrusion network-based big data. We proposed a combined Machine learning-based approach for the identification of pirated and malware files. First, the TensorFlow neural network is proposed to detect the pirated features of original software using software plagiarism. We collected 100 programmers' source codes files from KC99 to investigate the proposed approach.

The source code is pre-processed to clean from noise and to capture further the high quality features which include useful tokens. Then, TFIDF and LogTF weighting techniques are used to zoom the contribution of each feature in terms of source code similarity. The weighting values are then used as input to the designed Machine learning approach. Secondly, we proposed a novel methodology based on convolution neural network and colour image visualization to detect malware using Intrusion network.

## FUTURE SCOPE

**Advanced Threat Detection:**

Machine learning models will become more sophisticated, enabling the detection of previously unknown threats (zero-day attacks) by identifying patterns and anomalies that signify malicious behavior. Deep learning and reinforcement learning models will be utilized to improve the accuracy and speed of detection, reducing false positives and negatives.

**Real-time Analysis and Response**:

Future systems will employ real-time data analysis to detect and respond to attacks as they happen, minimizing damage and downtime. Integration of machine learning with automated response mechanisms will allow for immediate actions, such as isolating affected systems or blocking malicious traffic.

**Behavioral Analysis:**

Machine learning algorithms will focus more on behavioral analysis, learning the normal behavior of network users and systems to identify deviations that may indicate an attack.

User and Entity Behavior Analytics (UEBA) will become more prevalent, providing deeper insights into potential insider threats and compromised accounts.

**Integration with IoT and Edge Computing:**

As the Internet of Things (IoT) expands, machine learning models will be developed to secure IoT devices and networks, detecting vulnerabilities and potential breaches.

Edge computing will play a significant role, enabling machine learning models to process data locally on devices for faster and more efficient threat detection.

**References:**

1. Nutjahan, Farhana Nizam, Shudarshon Chaki, Shamim Al Mamun, M. Shamim Kaiser, "Attack Detection and Prevention in the Cyber Physical System",2016.
2. Yong Fang, Cheng Huang, Yijia Xu and Yang Li, "RLXSS: Optimizing XSS Detection Model to Defend Against Adversarial Attacks Based on Reinforcement Learning", 2019.
3. Rishikesh Mahajan, Irfan Siddavatam, "Phishing Website Detection using Ma- chine Learning Algorithms", 2018.
4. Vishnu. B. A, Ms. Jevitha. K. P., "Prediction of Cross-Site Scripting Attack Using Machine Learning Algorithms.", 2018.
5. Zohre Nasiri Zarandi,Iman Sharif, "Detection and Identification of Cyber-Attacks in Cyber-Physical Systems Based on Machine Learning Methods", 2020.