Research Article

A ROBUST CYBER SECURITY THREAT DETECTION MODEL USING ARTIFICIAL INTELLIGENCE TECHNOLOGY

Dr. V. Nagagopiraju[1] Panguluri Ashok[2] Kancheti Dhana Lakshmi[3] Chowdam Likhitha[4] Mandalapu Venkata Sasi Kumar[5]

[1]Department Of Cse & Ai, Chalapathi Institute of Engineering and Technology, Lam, Guntur, Andhra Pradesh, India.
[2]Department Of Cse & Ai, Chalapathi Institute of Engineering and Technology, Lam, Guntur, Andhra Pradesh, India.
[3]Department Of Cse & Ai, Chalapathi Institute of Engineering and Technology, Lam, Guntur, Andhra Pradesh, India.
[4]Department Of Cse & Ai, Chalapathi Institute of Engineering and Technology, Lam, Guntur, Andhra Pradesh, India.
[5]Department Of Cse & Ai, Chalapathi Institute of Engineering and Technology, Lam, Guntur, Andhra Pradesh, India

Abstract: The difficulty of ensuring cyber-security is steadily growing as a result of the alarming development in computer connectivity and the sizeable number of applications associated to computers in recent years. The system also requires robust defines against the growing number of cyber threats. As a result, a possible role for cyber-security might be performed by developing intrusion detection systems (ids) to detect inconsistencies and threats in computer networks.

An effective data-driven intrusion detection system has been created with the use of artificial intelligence, particularly machine learning techniques. This research proposes a novel twin support vector machine (tsvm) based security model which first considers the security features ranking according to their relevance before developing an ids model based on the significant features that have been selected. By lowering the feature dimensions, this approach not only improves predictive performance for unidentified tests but also lowers the model's computational expense. Trials are conducted using four common ml techniques to compare the results to those of the current approaches (decision tree, random decision forest, random tree, and artificial neural network). The experimental findings of this study confirm that the suggested methods may be used as learning-based models for network intrusion detection and demonstrate that, when used in the real world, they outperform conventional ml techniques.

Key words: Cyber Security, Cyber Security Threats, Intrusion Detection, Twin Svm

1. Introduction

Several new uses for computer and network technology have emerged in recent years, such as utilizing private data, public data, or commercial data. To stop system infiltration, cyber security has grown increasingly crucial. In the past, configuring a security policy on a firewall may not have provided adequate defence against such incursions due to the development of new types of invasions that make advantage of operating system flaws and message passing parameters, among other things. But, by employing the intrusion detection system (ids), it can both identify the issue and stop the incursion [1]. Cyber security has substantially improved in answer to the expanding range of cyber threats to prevent cybercrime and it's the term for a group of technologies, technical professionals, and procedures used to create security safeguards that keep cyberspace secure from cybercriminals. There are two primary methods of cyber security: automated cyber security and conventional cyber security. Many drawbacks of traditional cyber security, such as untrained individuals, inadequate system resource design, and restricted access to clean data, exacerbate cybercrimes [2].

The development of ai techniques has led to improvements in learning-based methods for spotting cyberattacks, and several studies have shown that they provide meaningful outcomes. To safeguard it systems against threats and suspicious network activity, however, is still very difficult due to cyberattacks' continual evolution. Effective defences and security concerns were given significant emphasis for creating trustworthy alternatives because of numerous network invasions and serious harm [3].

Along with this expansion, cyber criminals persisted in and broadened their fraudulent transactions, taking advantage of fresh security flaws and evading security measures to gain entry (hacking) to secret communication networks and inflict harm ranging from service interruption to the electronic theft of confidential or sensitive data or financial assets. The exponentially growing rate noted in the web ecosystem is mirrored by the growth rate of cyber-attacks, particularly those classified as aiassisted hacking, and has only recently begun to present a new challenge to the overburdened online security procedures, including many that require costly human analysis [4].

2. Literature Survey

The study discussed current deep learning (dl) techniques, conventional machine learning (ml) techniques, and ongoing research on using ai to prevent cybercrime. Lastly, evaluate the counterattacks that ai itself could face, identify the relevant defensive tactics, and study their characteristics. Unlike the natural intelligence that humans possess or the sort of intelligence that can be created in machines that operate and act like people, ai is the form of intellect that robots can display. Ai also known as ml, is essentially the term used when a computer acts like a person while doing tasks like problem-solving or learning. The article presented the ai methods may be used to cybersecurity and highlight a few of the intelligent based strategies that are currently in use. In addition, the shortcomings of ai-based cybersecurity techniques were emphasised and potential future research options were provided. A system that detects intrusions using data may be constructed using ai, especially ml methods.

3. System analysis

3.1 Existing System

The existing system for your project, "cyber security threat detection model using artificial intelligence technology," leverages machine learning techniques, particularly the binary grasshopper optimized twin support vector machine (bgotsvm) model. This system is designed to address the escalating challenge of cybersecurity by detecting inconsistencies and threats within computer networks. It begins with the collection and preprocessing of network data, followed by feature ranking to prioritize relevant security features. The ids model is then developed based on the selected significant features, reducing feature dimensions and computational costs. Comparative experiments are conducted against traditional machine learning approaches like decision tree, random decision forest, random tree, and artificial neural network. The system's experimental results demonstrate its superiority in network intrusion detection and its potential for real-world cybersecurity applications.

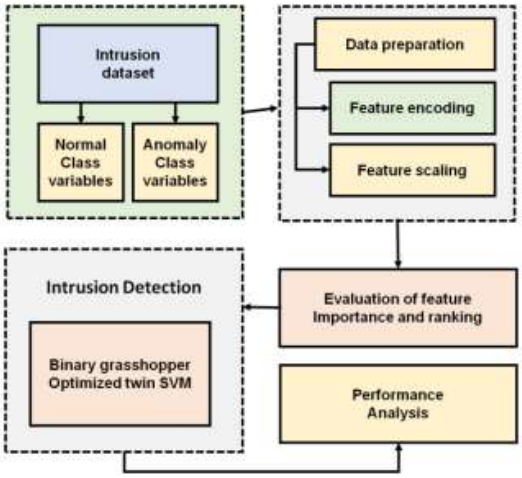
Limitations of existing system

Limited scalability: the system may face challenges in scaling to handle large and complex network environments, making it less suitable for organizations with extensive and diverse network infrastructures.
Dependency on feature selection: the effectiveness of the system relies heavily on the initial feature selection process, which may not always guarantee optimal feature relevance, leading to potential false positives or negatives.

3.2 proposed system

The proposed system for the project, "cyber security threat detection model using artificial intelligence technology," aims to address the limitations of the existing system. It seeks to enhance scalability by incorporating distributed computing and cloud-based solutions, making it suitable for large and complex network environments. The proposed system will focus on automating the feature selection process using advanced feature engineering techniques, reducing the reliance on manual feature selection. It will also implement real-time detection capabilities for immediate threat response. To improve interpretability, the system will incorporate explainable ai techniques, making its decisions more understandable for cybersecurity professionals. Additionally, the proposed system will emphasize adaptability and continuous learning, integrating threat intelligence feeds and ensuring timely updates to stay ahead of evolving cyber threats.

4. System architecture


Intrusion dataset
Normal Class variables
Anomaly Class variables
Data preparation
Feature encoding
Feature scaling
Intrusion Detection
Binary grasshopper Optimized twin SVM
Evaluation of feature Importance and ranking
Performance Analysis

5. Methodology

Information process involves feature encoding, scaling, and analysis of the intrusion dataset's properties. Encoding of features: the dataset contains the specified security features' numerical and fictitious values, as previously mentioned. In all generally valued characteristics must be transformed into the targeted ml-based intrusion detection models in order to match that data to that model's vectors. Although "one hot encoding" is a common method, in this study "label encoding" is used. The cause is because a large rise in the number of feature dimensions occurs in one hot encoding approach. The label-encoding method, on the other hand, immediately transforms the feature values into specific numeric values. Scaling features: data pre-processing and feature scaling are both terms used in data pre-processing. The security feature values are divided into a number of ranges that vary depending on the feature. The characteristic variables' range, sometimes referred to as the individual variables, is normalized using the data scaling approach. To achieve this, a conventional capability approach was used, which moderates the security mechanisms with a mean value of 0 and a standard deviation of 1. After standardization, the variables are ready for additional analysis to create the security model let $A \in Q\ n1\times m$ symbolize the pieces of data in class +1 and $B \in Q\ n2\times m$ reflect the pieces of data in class 1, with r standing for random numbers. Every entry in both matrix correlates to a test dataset. The nonparallel hyper - parameters are found via the linear twin support vector machine (tsvm) such that every hyperplane is closest to and furthest from the pieces of data of one class, respectively. It is given the class name +1 or 1 depending on how close the data samples from the two hyperplanes are to one another. The two quadratic programming problems (qpps) are solved by tsvm using the optimal solution of one class and constraints related to the other class.

6. Modules

The modules for the "cyber security threat detection model using artificial intelligence technology" project can be outlined as follows:

Data collection and pre-processing: collect and pre-process network data, logs, and activity parameters for analysis, ensuring data quality and consistency.

Feature engineering and selection: implement advanced feature engineering techniques to extract relevant security features and automate feature selection processes to prioritize critical attributes.

Model development: develop the core intrusion detection system (ids) using the binary grasshopper optimized twin support vector machine (bgotsvm) model, optimizing the chosen algorithm for enhanced accuracy.

Real-time monitoring and alerting: integrate real-time monitoring capabilities to continuously analyze network data and trigger immediate alerts upon detecting suspicious or malicious activities.

Explainability and adaptation: incorporate explainable ai techniques for improved model interpretability and implement mechanisms for continuous adaptation, integrating threat intelligence feeds and timely updates to stay current with evolving cyber threats.

7. Result

8. Conclusion

It professionals, e-commerce professionals, and application developers are extremely worried about the practicality and potential of an ml-based intrusion detection modelling for security reasons. Several cyber attack types with relevant features are often included in a cyber-security data collection. Thus, certain classifiers that are based on a variety of assault kinds and a number of different factors could not function well in terms of accuracy and true prediction rate. ➢ the effectiveness of the bgotsvm model in this work has been discussed. Recall, f1-score, and overall accuracy are only a few of the performance measures that have been evaluated.

Future scope: in order to offer the community of cyber-security experts automated security services in the future, it is intended to increase the datasets related to cyber-security and create a data-driven intrusion detection system.

References

1) p.sornsuwit, and s.jaiyen, "a new hybrid machine learning for cybersecurity threat detection based on adaptive boosting," applied artificial intelligence, 33(5), pp.462- 482, 2019.
2) k.shaukat, s.luo, s.chen, and d. Liu, "cyber threat detection using machine learning techniques: a performance evaluation perspective,"in ieee international conference on cyber warfare and security. Ieee, october2020,pp. 1-6.
3) q.h. vu, d.ruta, and l.cen, "gradient boosting decision trees for cyber security threats detection based on network events logs," in 2019ieee international conference on big data (big data). Ieee, december2019, pp. 5921-5928.
4) j. Lee, j. Kim, i.kim, and k. Han,"cyber threat detection based on artificial neural networks using event profiles,"ieee access, vol. 7, pp.165607-165626, 2019.
5) j.h. li, "cyber security meets artificial intelligence: a survey;" frontiers of information technology & electronic engineering, vol. 19, no.12, pp.1462-1474, 2018.
6) n. Rawindaran, a.jayal, e.prakash, and c.hewage, "cost benefits of using machine learning features in nids for cyber security in uk small medium enterprises (sme)," future internet, vol. 13, no. 8, p.186, 2021.
7) r.prasad, v.rohokale, r.prasad, and v.rohokale, "artificial intelligence and machine learning in cyber security," cyber security: the lifeline of information and communication technology, pp.231-247, 2020
. 8) t.c.truong, i.zelinka, j.plucar,m.čandík, and v.šulc,"artificial intelligence and cybersecurity: past, presence, and future," in artificial intelligence and evolutionary computations in engineering systems, pp. 351-363, springer singapore, 2020.
9) i.h.sarker, y.b.abushark, f.alsolami, and a.i. khan,"intrudtree: a machine learning based cyber security intrusion detection model," symmetry, vol. 12, no. 5, p.754, 2020.
10) diro, and n.chilamkurti, "distributed attack detection scheme using deep learning approach for internet of things," future generation computer systems, vol. 82, pp.761-768, 2018.