Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 2, pp. 760–769, DOI 10.12694/scpe.v25i2.2627

# DETECTION AND PREVENTION OF CYBER DEFENSE ATTACKS USING MACHINE LEARNING ALGORITHMS

YONGQIANG SHANG*

**Abstract.** Recent advancements in computing power, memory capacities, and connectivity have led to a corresponding surge in the utilization of big data, online platforms' prevalence, and machine learning's sophistication. Concerns regarding safety and the need for state-of-the-art security tools and methods to counter evolving cybercrime are amplified by the swift digitization of the world. This study investigates defensive and offensive applications of machine learning in cybersecurity. Additionally, it explores potential strategies to mitigate cyberattacks on machine learning models. The focus is on how machine learning facilitates cyberattacks, including developing intelligent botnets, advanced phishing using spear techniques, and deploying stealthy malware. Furthermore, the paper highlights the significance of artificial intelligence in digital safety, emphasizing its role in malware analysis, network vulnerability assessment, and threat prediction.

**Key words:** Internet risk estimation, Machine learning, Threat detection, Cyber safety, App-level protection

**1. Introduction.** Cryptography remains a top priority as the world transitions to a digital infrastructure. With advancements in network infrastructure, such as the World Wide Web, gaining access to cutting-edge science and technological breakthroughs has never been more convenient. Freeware and academic articles are increasingly accessible to the general public online. However, security experts and malicious players have easy access to state-of-the-art technology and research, each with their motivations for manipulating them. Safety measures have benefited from the development of techniques and innovations made possible by studies and advances in artificial intelligence. These technologies simplify the identification and effective counteraction of potential security threats. But, hackers can leverage this data to plan and execute more sophisticated and widespread attacks. Hackers hold a significant advantage in the ongoing battle on the internet, as they must only be successful once out of numerous attempts [1].

Achieving a 100% success rate is imperative when the goal is safety. Numerous studies revealed that in 2017, hackers targeted a diverse array of businesses, individuals, and mobile applications. Documents such as surveillance records, accounting information, and personal data were among the stolen information. The potential consequences of this data falling into the wrong hands, whether the general population or the black market, could be catastrophic. Here are some research findings on the impact of cybersecurity on businesses, organizations, and individuals [2]:

- Recently, over 350 trillion dollars have been lost or stolen due to cybercrime, including the costs incurred for repairing damages caused by criminal activities.
- A shortfall of over 1.8 million cybersecurity professionals is projected by 2022.
- Companies must spend at least $100 billion annually to keep up with evolving technology.
- Intruders generate over a billion dollars yearly through ransomware attacks like WannaCry and CryptoWall.

Keeping step with and combating the escalating complexity of intrusions is becoming increasingly challenging, given the rapid obsolescence of safeguards. The average time to identify an intrusion is approaching seventy days. The growing magnitude and intricacy of intrusions further complicate efforts to keep up with the continuously emerging new threats and vulnerabilities. This challenge arises from the increasing sophistication and breadth of intrusions. Here, the authors focus on artificial intelligence, an emerging discipline with far-reaching implications for online safety [3].

*Xinyang Agriculture and Forestry University, Department of Information Engineering, Xinyang Hennan 464000, China (yongqiangshang8@163.com)

**1.1. Defence against cyber-attacks using machine learning .** Machine learning is a subset of AI that enables machines to learn and improve from data without being programmed. It employs mathematical frameworks derived from the analysis of patterns in datasets to achieve this objective. These models are then applied to make predictions based on newly acquired information. For example, machine learning systems are utilized in the e-commerce sector to customize product recommendations for individual shoppers. Moreover, machine learning is finding applications in the medical field for predicting outbreaks and assessing whether a patient is predisposed to developing cancer based on their medical history. Machine learning holds considerable potential for diverse applications across various fields [4].

Machine learning techniques are broadly categorized into prospective (controlled) artificial intelligence methods and pattern identification (or uncontrolled) machine learning techniques. In controlled learning, an algorithm is provided with a set of training data and tasked with predicting the value of a target variable using various learning methods. For example, a machine learning model may determine whether an internet protocol (IP) address is involved in a distributed denial of service (DDOS) assault based on factors such as the geographical origin of the IP, the frequency of web requests, and the time of day the requests were made. Various machine learning methods, including linear and logistic regression, decision trees, and support vector machines (SVMs), fall under the umbrella of supervised learning. Conversely, the objective of unsupervised learning is not to predict a specific variable. Uncontrolled algorithms learn to uncover relevant patterns and correlations in datasets. For instance, clustering and association algorithms may be employed to identify groups of malicious software with common operations and behaviour [5].

The domain of cybercrime and security is rapidly adopting machine learning. Beyond being a prevalent application area, the IT sector offers various potential uses, including malware analysis and log analysis, where machine learning is widely deployed. Both well-intentioned users and those with malicious intent leverage the powerful machine-learning capabilities of the internet. The upcoming section will explore the dual nature of machine learning's application in internet security and criminal activities [6].

**1.2. Machine Learning Techniques for Detecting Cyber Attacks.** Two approaches to cyber threat detection are signature and anomaly-based, employing ML techniques. Recent advancements in artificial intelligence methods have facilitated the establishment of signs to identify the code and behaviour of spyware accurately. These signatures are integrated into the construction process. Several strategies have been developed for the swift and efficient retrieval of signatures from versatile parasites, including NSG and LSEG, which are methods for creating signatures based on network behaviour. The LESG method primarily focuses on viruses propagated through buffer overflow attacks. Figure 1 illustrates the attack vectors in the world [7].

The F-Sign method extracts a signature from the worm's code, which can be utilized to detect and halt the worm's propagation. Alternative methods for generating signatures of malicious software based on the network traffic it produces include semantic aware (SA), as documented in scholarly publications. These technologies can accurately detect malicious activity even when the system has considerable noise. Building a model using anomaly-based approaches is a common practice for identifying cyberattacks, encompassing typical and typical network activity patterns [8]. These methods commonly employ unsupervised, semi-supervised, and supervised algorithms, all rooted in artificial intelligence. Techniques like k-means, fuzzy c-means, QT and SVM are frequently utilized for constructing clustering approaches in unsupervised learning. When these techniques aggregate network activity, a decision often arises regarding whether or not the cluster should be labelled malicious. Completely unsupervised algorithms adhere to a common rule, positing that the most significant clusters are the only consistent ones. As a result, routine network occurrences exhibit no signs of an attack. In reality, individuals must determine which combinations should be deemed abnormal. Supervised machine-learning strategies require at least one learning stage to be completed before initiating the traffic model construction [9].

In most cases, learning concludes once internet traffic patterns are sufficiently processed. Numerous controlled anomaly-based systems for detecting network intrusions employ various machine-learning methods. The proposed approaches typically divide the attack detection process into two phases: the first, termed feature vector extraction, and the second, known as the method development phase. For instance, developers utilized information theory to identify cyber threats, employing a metric based on entropy and information gain. Outliers were identified using a linear classification algorithm [10].

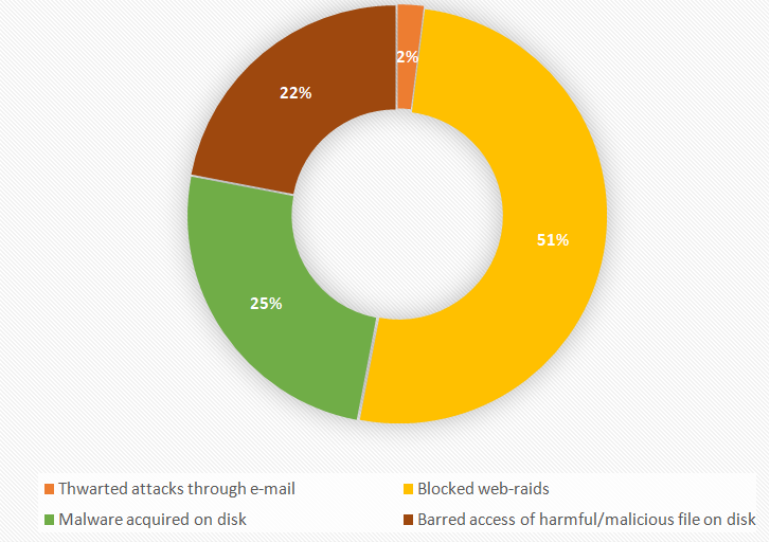


Fig. 1.1: The attack vectors worldwide in the year 2022

Unusual App activity is detected using a k-NN classifier with OS events, such as the quantity of started processes and system interactions. The authors applied a k-NN classification with metrics to the KDD sample to detect SYN Flooding, U2R, and remote-to-local (R2L) attacks. Classifiers and neural networks were jointly employed for spam detection. To identify DoS assaults, the researchers utilized a statistical approach. Feature vectors were used to train the Naive Bayes predictor, encompassing various UDP and TCP packets and their sizes. Attributes were successfully calculated using discrete wavelet transform (DWT) and combining retrieval based on a broad range of network properties. More sophisticated techniques, such as Hidden Markov Models, were used to identify denial-of-service and application layer assaults. Additionally, artificial neural networks are increasingly used in intrusion detection [11].

RBF artificial neural networks are applied to analyze network data for anomalies. Artificial neural networks prove effective in identifying UDP flooding incidents. Internet surveillance techniques based on support vector machines have also been investigated. The existing literature may offer solutions utilizing rough set theory and semi-supervised learning. The genetic method was employed to detect outliers. After the learning process, a set of rules is generated to characterize normal and aberrant traffic patterns in the network. The program underwent testing using data from the DARPA database. The feature vector encompassed the TCP connection's length, the amount of data transferred, and the link's originating and final IP addresses. The evolutionary algorithm and the correlation method were utilized to identify SQL Injection threats [12].

The subsequent sections of the article are organized as follows. The related works provide an overview of artificial intelligence, addressing supervised and unsupervised learning. The third section examines the proposed methodology for online safety, including network risk assessment and identifying and preventing infections. The fourth section explores the utilization of machine learning in cybercrime, with its simulated experiments and results.

**2. Related works.** Detecting intrusions and other abnormalities within the underlying Internet of Things (IoT) network is becoming an increasingly urgent concern. As IoT technology is deployed across various industries, the number of attacks and threats against it has also increased. Attack disruption attacks, data type searching, fraudulent authority, fraudulent functioning, scanning, espionage, and improper configuration are just a few of the attacks and irregularities that can compromise an IoT network. In this analysis, we compare the results of various machine learning models to assess their ability to predict incidents and anomalies on IoT devices reliably. Machine learning (ML) methods such as logistic reconstruction, SVM, RF, decision trees, and neural network training have all found applicability in this context. Performance levels are compared and

contrasted using criteria for precision, recall, reliability, F1 scores, and the area under the receiver operating characteristic curve [13].

Random Forest demonstrates superior performance across various metrics, even when the accuracy of the two approaches is comparable. In particular, the RF technique outperformed other methods' ability to detect and clear assaults effectively. The comprehensive study concluded that, based on the dataset analyzed, the RF approach is recommended for addressing intrusion issues in IoT networks, leading to these findings. It provided more accurate sample-level predictions for denial of service (DoS) and normal instances than any previous approach. These results suggest that RF analysis is well-suited for this investigation [14].

This study employs traditional ML methods on the dataset and then compares them without introducing any novel methodologies based on this dataset. Consequently, there is a need for further research to develop a reliable detection method, including a deeper exploration of the framework's general structure. The data from a simulated setting presents potential challenges when transitioning to real-time data applications. Addressing these challenges requires a more empirical study focusing on real-time data to provide a more comprehensive solution. The varying patterns of behaviour exhibited by different micro-services in the network contribute to the irregular behaviour of IoT services at different times. More research is necessary to understand these dynamics thoroughly. Despite achieving an impressive 99.4 per cent accuracy, the study's findings indicate that while RF performs well compared to other approaches, its continued effectiveness is not guaranteed with large datasets or unexpected scenarios. Therefore, additional research is suggested [15].

The hardware and computer software can be the foundation for an intrusion detection system. The primary objective is to monitor a system or a system of systems to ensure that no malicious or policy-breaking behaviour goes undetected. Intrusion detection systems (IDS) consider various network-related variables, such as source addresses, procedures, and banners, to assess the irregularity of behaviour. The fundamental goal of any security detection system is to achieve the highest level of precision while minimizing false alarms [16].

In intrusion detection, the study aims to identify the essential building blocks required for creating a detection framework. The model utilizes an ensemble method for identification that requires significantly less processing power, addressing challenges faced by conventional group-based surveillance algorithms. The article employs the Chi-square feature selection and a classification array consisting of SVM, modified naive Bayes (MNB), and LPBoost to construct an intrusion detection model. Chi-square feature selection analyzes variance to control the importance of each feature, focusing on those that matter in determining a class. Laboratory findings demonstrate the LPBoost ensemble's superior precision compared to baseline classifications. The most effective method for predicting the class label involves using a majority voting system, such as support vector machines, multi-network averages, or LPBoost, instead of a single classification. Given the substantial class disparity throughout the network traffic, this is preferable, providing a more robust approach than a single classification [17].

In the contemporary era marked by a digital revolution, knowledge storage, accessibility, and dissemination over the web have experienced significant and exponential growth. The emergence of innovations based on the IoT has further contributed to eliminating digital barriers and enhancing data and information exchange across pervasive networks. These IoT innovations have notably improved communication between devices, leading to heightened concerns among consumers regarding data theft, privacy breaches, and the secure transmission of their data and information over the web [18].

Various approaches, including deploying detection and prevention devices, are employed to combat data theft and other online data safety risks. This research specifically compares and contrasts two types of intrusion detection systems. The first type utilizes a machine learning strategy called SVM, while the second type employs an association rule data mining methodology known as Apriori. Both methods have originated from academic research. Research findings indicate that SVM surpasses Apriori in terms of accuracy, while Apriori demonstrates greater efficiency in testing duration [19].

Over the past decade, there has been a significant increase in individuals relying on the internet as their primary information resource. Technological advancements such as the IoT and high-performance computing (HPC) have accelerated the pace of online data access and the sheer volume of information generated. This surge in data production and accessibility has made safeguarding sensitive information from unauthorized access more challenging. Consequently, developing analytics and innovative models to mitigate threats posed by attackers

and scammers to private data stored online has become imperative [20].

The approach to enhance security is to limit the quantity of data stored online. In this study, the authors evaluated and compared the performance of two feature and two categorization methods for intrusion detection, aiming to mitigate the impact of data theft and other real-time network security breaches on computer systems. Leveraging the gathered information, researchers developed a robust intrusion model to prevent real-world attempts at network penetration in live software settings. Cyber-physical systems (CPS) integrate digital information processing capabilities and communication networks with tangible components and operations. Among the various anomalies that can impact the proper functioning of these systems, safety breaches and breakdowns are particularly common. While extensive research has focused on defect diagnosis and security analysis in CPS independently, the crucial and timely challenge of distinguishing between different sources of irregularities, such as flaws and attacks, remains inadequately addressed [21].

This study concentrates on the energy-aware smart home (EASH) system and its internal communication ecosystem. Specifically, the authors define the challenge of distinguishing between component breakdowns and network attacks in EASH, considering how these elements influence information exchange patterns. Key contributions of this paper include a formal demonstration of the relationship between these irregular sources and an ML-based methodology for identifying the issue. The developed system undergoes testing in both simulated and real-world laboratory environments, with findings indicating the potential to achieve a classification accuracy rate exceeding 85% [22].

Analysis of the classes and features employed in the proposed method suggests significant classification precision improvements based on our laboratory findings. The results from these tests substantiate these hypotheses. The study aims to explain the outcomes of a task to differentiate between anomalies affecting an EASH system. This suggests that observed anomalies could stem from malfunctioning hardware or malicious network activity. A discrimination operator was developed after studying the connection between various abnormalities and their impacts on the system's information channels [23].

Utilizing ML-based categorization algorithms sets the approach apart. This research concludes that supervised machine learning methods offer a viable option for distinguishing between faulty and attack classes with high reliability. Explorations into misclassifications of cases with similar effects on the network were conducted in simulation and real-time testbed experimental environments, encompassing analyses of aberrant classes and the considered properties. The findings suggest modifying description datasets by adding or removing features could enhance categorization outcomes [24].

The widespread adoption of IoT-connected devices is rapidly approaching, with IoT services expanding to achieve ubiquitous integration. The increasing popularity of IoT-enabled gadgets and services has led to many risks and attacks against them. While intrusions on the IoT are not new, the growing embedding of IoT in our lives and society underscores the critical need to enhance cybersecurity measures. Understanding potential threats and attacks against IoT infrastructure has become imperative. The research aims to explore the diverse dangers that could impact IoT devices and offerings, analyzing and characterizing the incursions and attacks directed at them. Recognizing potential threats is a prerequisite for implementing effective protective measures for the IoT. This article focuses on cybersecurity threats to the IoT, with the primary objective being identifying resources and inventory of potential threats, vulnerabilities, and attacks against the IoT [25].

A presentation highlighted security vulnerabilities associated with IoT devices and services, summarizing the most pressing issues affecting IoT safety. Secrecy, anonymity, and trust in entities were identified as challenging security aspects. The study emphasized that addressing privacy and security concerns is crucial for ensuring the security and user-friendliness of IoT devices and services. The article also investigated cyber threats driven by the unique characteristics of cyberspace, including individuals, motives, and capabilities. Evidence suggests that threats from government intelligence services and organized criminal groups pose more formidable challenges than lone hackers, as they may target fewer predicted goals, with a single attack expected to have a more significant impact [26].

The study concluded that manufacturers and consumers are crucial in ensuring IoT security. To address the weaknesses in the current security infrastructure, new standards for the IoT are essential. This research aims to contribute to future endeavours to enhance the understanding of vulnerabilities in IoT infrastructure and assess the likelihood of attacks against IoT and the consequences of such assaults. Early considerations in

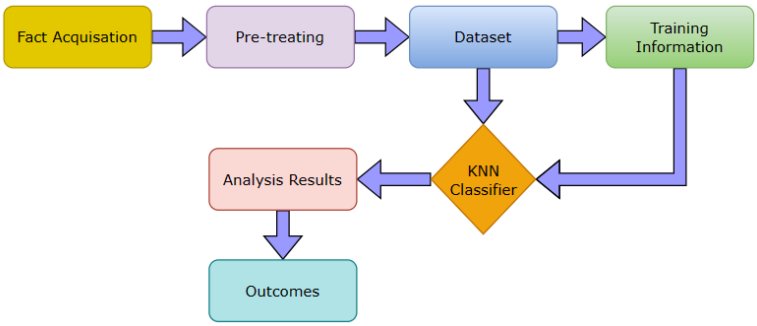


Fig. 3.1: The workflow of the proposed technique

the product creation process should include security measures such as access control, authentication, identity management, and a flexible reputation management system. This survey is anticipated to benefit security experts as it illuminates the most critical issues related to IoT security, offering a deeper understanding of the threats and their features [27].

**3. Proposed Methodology.** In this investigation, we employ an approach to intrusion detection and propose measures to enhance the security of innovative environments. The diagram illustrates the utilization of a K-nearest neighbours (KNN) classifier in the proposed method for intrusion identification. Following the dataset selection, a preliminary processing phase is implemented, converting certain symbols to numerical values to enhance the accuracy of the detection rate. The subsequent training phase involves the application of the KNN technique, training the model using 65% of the dataset. The final step involves evaluation, wherein the remaining 35% of the data assesses the model's accuracy. To further measure the performance of the proposed system, metrics such as the accuracy detection rate and the occurrence of four types of warnings are estimated in this phase. The architecture of the proposed model is explained in Figure 3.1.

Following the collection of information, preparation becomes necessary due to the unstructured nature of the features, which consist of text or numeric values. K-nearest neighbours (KNN) specifically deal with numerical data during the training and testing. Therefore, the objective of the preliminary processing stage is to convert symbols and characters into their numerical equivalents. The training and evaluation stages are crucial in developing an intelligent model. During the training phase, KNN enhances reliability, reduces false alarms, and assesses the false alarm rate. The dataset is then partitioned into a training set (comprising 60% of the total) and a testing set (representing 40%). Testing holds significance as it determines the model's effectiveness, with metrics such as precision, recall, and F1-score being computed at this stage.

**3.1. Dataset Description.** To establish an effective detection and prevention framework, having an accurate and current database is essential for concluding the operations of various cyberattacks. In this context, we leverage the multi-step cyberattack dataset (MSCAD), which comprises six individual files such as MSCAD.xlsx, N0, Scan, App01, App02, WB01, and WB02.

The annotated form of the provided dataset can be located in the MSCAD.xlsx file. Each of the six provided PCAP documents underwent analysis in OpenVPN. The timestamp is crucial in assessing whether network communication is malicious or benign. After processing these PCAP documents, the MSCAD dataset comprised 79 features with corresponding labels. The primary attack concept in MSCAD is rooted in a password-breaking model. MSCAD's secondary attack methodology predates the volume-based distributed DoS concept.

**3.2. K-Nearest Neighbors.** K-nearest neighbours (KNN) can effectively address regression and classification problems, offering simplicity in theory and practical implementation. However, its efficiency tends to decrease as the volume of data increases, resulting in sluggish performance. The training step, on the other hand, is highly time-efficient, as the data itself serves as a model for future identification, enabling spontaneous and randomized programming based on the provided information.

**3.3. Evaluation Metrics.** Assessment metrics are crucial in testing and refining classification algorithms within an intelligent intrusion detection system (IDS). These metrics are employed to evaluate the effectiveness of the IDS. In this study, precision is utilized to assess the effectiveness of KNN. Using Equation 3.1, the precision is defined as the percentage of detected activities correctly categorized [28].

$$Precision = \frac{\text{actions accurately categorised}}{\text{sum total of actions}} \tag{3.1}$$

The efficiency of a system can be quantified by calculating the occurrence of four distinct sorts of alarms (the ambiguity matrix). All four possible outcomes are true positive (TP), false positive (FP), true negative (TN), and false negative (FN) are identical.

$$TP = \frac{TP}{TP} + FN \tag{3.2}$$

$$TN = \frac{TN}{TN} + TP \tag{3.3}$$

$$FN = \frac{FN}{FN} + TP \tag{3.4}$$

$$FP = \frac{FP}{FP} + TN \tag{3.5}$$

In addition, Equation 3.6 and 3.7 are used to calculate accuracy and recall.

$$Accuracy = \frac{TP}{TP} + FN \tag{3.6}$$

$$Recall = \frac{TP}{TP} + FP \tag{3.7}$$

The F1-score is also used as a periodic mean of both recall and accuracy. The formula for the F1 score is as follows:

$$F1 - score = 2 \times (Precision \times \frac{Recall}{Precision} + Recall) \tag{3.8}$$

The relationship between the model size and recall, serving as a comprehensive guide that reveals the connections and dependencies within the framework, is represented in Figure 3.2.

The efficiency anomalies concentrating on the effects of employing distinct standard frameworks for each individual and URL response are shown in Figure 3.3. This figure is a pivotal reference point, shedding light on the complexities of using separate frameworks.

**4. Experimentation & Results.** This research divides the dataset into a training set and an evaluation set at a ratio of 56% to 39%. The proposed intrusion detection system is then assessed based on accuracy, precision, recall, and F1 score. The experiments are conducted on a computer with 64 GB of DDR4 memory, two gigabytes of dedicated graphics processing RAM, and two Intel Xeon 8 Core processors running at 2.43 GHz, housed in Dell PowerEdge T430 servers.

As indicated in Table 4.1, the model achieves an accuracy rate of 82.75%. This implies that the outcomes obtained using KNN machine-learning technique for detecting various attacks are generally reliable. To measure the system's capability in identifying intrusive situations, we utilize the MSCAD dataset, as demonstrated in the evaluation of the datasets. Figure 4.1 provides a comparison of our findings with the work of others.

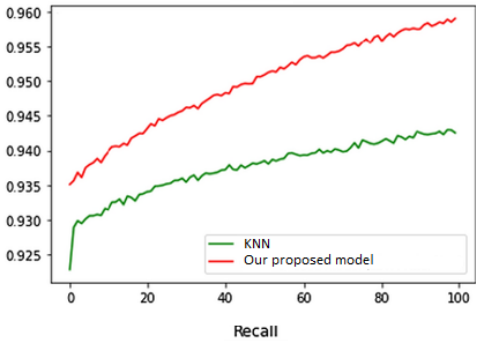


Fig. 3.2: The recall versus model size

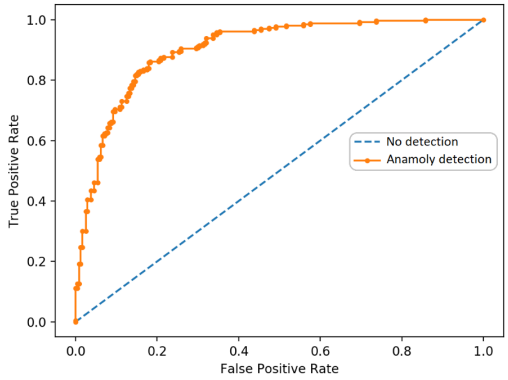


Fig. 3.3: Performance of finding anomalies using a separate normal model for every URL request

Figure 4.1 displays the accuracy values for three models: an IoT-based model with 94.6%, an ML-based model with 96.2%, and a proposed model with 98.5%. These accuracy percentages signify the proportion of accurate predictions made by each model, with the proposed model achieving the highest accuracy, followed by the ML-based model and the IoT-based model. The findings suggest that the proposed model performs exceptionally well, showcasing a strong ability to make precise predictions.

The model's accuracy rate of 93.10% indicates the overall correctness of its predictions, showcasing the proportion of accurately classified instances in Table 4.2. The recall rate, at 90.22%, measures the model's efficacy in correctly identifying relevant instances within the dataset. The F1-Score, reaching 91.42%, provides a balanced assessment by considering both precision and recall. Together, these metrics furnish valuable insights into the model's multilayered performance, contributing to its predictive capabilities across diverse dimensions.

**5. Conclusion.** In this article, the application of AI is explored in security, considering both protective and threatening perspectives, and has discussed potential challenges faced by machine learning-based models. Machine learning offers an efficient approach to automating complex online operations for offensive and preventive purposes. As machine learning becomes an integral part of hackers' cyber cache, it is anticipated that the complexity and diversity of AI-based attacks will increase. Computer education and security professionals are advised to stay updated on the modern developments in ML, including adversarial learning, to leverage this knowledge for national security. This study sets the stage for further research into the challenges of implementing scalable cybercrime systems using artificial intelligence in diverse operational contexts. This research focuses on detecting cybersecurity issues using a machine learning-based approach, specifically employing the KNN technique. The proposed system's evaluation involves TP, TN, FP, FN, and measures such as F1-score, accuracy,

Table 4.1: Detection accuracy rates

| Alarms style | Accuracy Rate (%) |
|---|---|
| TP | 94.32 |
| TN | 91.97 |
| FP | 5.68 |
| FN | 8.03 |
| Precision | 82.75 |

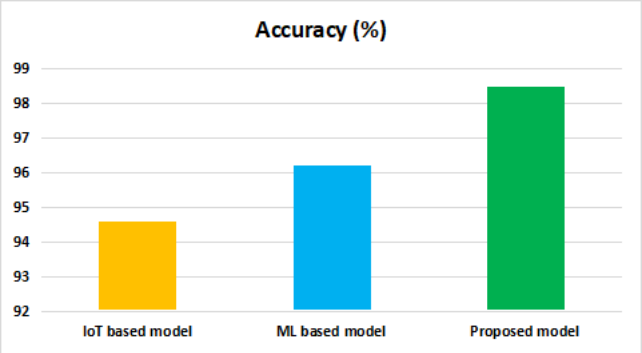


Fig. 4.1: Accuracy Rate Comparison

precision, and recall are calculated, revealing promising results.

**6. Acknowledgement.** The study was supported by Key R&D and Promotion Special Project (Science and Technology Research) in Henan Province (232102210146).

REFERENCES

[1] B. Ahmad, W. Jian, and Z. Anwar Ali, "Role of Machine Learning and Data Mining in Internet Security: Standing State with Future Directions," J. Comput. Networks Commun., vol. 2018, 2018, doi: 10.1155/2018/6383145.
[2] Li, L., He, W., Xu, L., Ash, I., Anwar, M., & Yuan, X. (2019). Investigating the impact of cybersecurity policy awareness on employees' cybersecurity behavior. International Journal of Information Management, 45, 13-24.
[3] Vocke, C., Constantinescu, C., & Popescu, D. (2019). Application potentials of artificial intelligence for the design of innovation processes. Procedia CIRP, 84, 810-813.
[4] S. Dolev and S. Lodha, "Cyber Security Cryptography and Machine Learning ", In Proceedings of the First International Conference, CSCML 2017, Beer-Sheva, Israel, June 29-30, 2017.
[5] Fenil, E., & Mohan Kumar, P. (2020). Survey on DDoS defense mechanisms. Concurrency and Computation: Practice and Experience, 32(4), e5114.
[6] Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., Al-Nemrat, A., & Venkatraman, S. (2019). Deep learning approach for intelligent intrusion detection system. Ieee Access, 7, 41525-41550.
[7] Bagui, S., Kalaimannan, E., Bagui, S., Nandi, D., & Pinto, A. (2019). Using machine learning techniques to identify rare cyber-attacks on the UNSW-NB15 dataset. Security and Privacy, 2(6), e91.
[8] AlEroud, A., & Alsmadi, I. (2017). Identifying cyber-attacks on software defined networks: An inference-based intrusion detection approach. Journal of Network and Computer Applications, 80, 152-164.
[9] Aboueata, N., Alrasbi, S., Erbad, A., Kassler, A., & Bhamare, D. (2019, July). Supervised machine learning techniques for efficient network intrusion detection. In 2019 28th International Conference on Computer Communication and Networks (ICCCN) (pp. 1-8). IEEE.
[10] Gao, Y., Li, X., Peng, H., Fang, B., & Philip, S. Y. (2020). Hincti: A cyber threat intelligence modeling and identification system based on heterogeneous information network. IEEE Transactions on Knowledge and Data Engineering, 34(2), 708-722.
[11] Seo, J. H., & Kim, Y. H. (2018). Machine-learning approach to optimize smote ratio in class imbalance dataset for intrusion detection. Computational intelligence and neuroscience, 2018.

Table 4.2: Performance Metrics for Model Evaluation

| Estimation metrics | Rate (%) |
|---|---|
| Accuracy | 93.10 |
| Recall | 90.22 |
| F1-Score | 91.42 |

[12] Jemal, I., Cheikhrouhou, O., Hamam, H., & Mahfoudhi, A. (2020). Sql injection attack detection and prevention techniques using machine learning. International Journal of Applied Engineering Research, 15(6), 569-580.
[13] Alsouda, Y., Pllana, S., & Kurti, A. (2019, May). Iot-based urban noise identification using machine learning: performance of SVM, KNN, bagging, and random forest. In Proceedings of the international conference on omni-layer intelligent systems (pp. 62-67).
[14] R. Primartha and B. A. Tama, "Anomaly detection using random forest: A performance revisited," Proc. 2017 Int. Conf. Data Softw. Eng. ICoDSE 2017, vol. 2018-Janua, pp. 1–6, 2018, doi: 10.1109/ICODSE.2017.8285847.
[15] M. Hasan, M. M. Islam, M. I. I. Zarif, and M. M. A. Hashem, "Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches," Internet of Things (Netherlands), vol. 7, p. 100059, 2019, doi: 10.1016/j.iot.2019.100059.
[16] Y. Cheng, M. Naslund, G. Selander, and E. Fogelström, "Privacy in machine-to-machine communications A state-of-the-art survey," International Conference on Communication Systems (ICCS), pp. 75-79, 2012.
[17] Jiang, K., Wang, W., Wang, A., & Wu, H. (2020). Network intrusion detection combined hybrid sampling with deep hierarchical network. IEEE access, 8, 32464-32476.
[18] Porkodi, R., & Bhuvaneswari, V. (2014, March). The internet of things (IOT) applications and communication enabling technology standards: An overview. In 2014 International conference on intelligent computing applications (pp. 324-329). IEEE.
[19] De Almeida, M. B., de Pádua Braga, A., & Braga, J. P. (2000, November). SVM-KM: speeding SVMs learning with a priori cluster selection and k-means. In Proceedings. Vol. 1. Sixth Brazilian Symposium on Neural Networks (pp. 162-167). IEEE.
[20] Nan, A. A., Shawky, M. M., Ahmed, A. M., & Ellaithy, D. M. (2021, December). Design and Implementation of High-Performance Computing Unit for Internet of Things (IoT) Applications. In 2021 International Conference on Microelectronics (ICM) (pp. 258-261). IEEE.
[21] Monostori, L., Kádár, B., Bauernhansl, T., Kondoh, S., Kumara, S., Reinhart, G., & Ueda, K. (2016). Cyber-physical systems in manufacturing. Cirp Annals, 65(2), 621-641.
[22] Kabir, M. H., Hoque, M. R., & Yang, S. H. (2015). Development of a smart home context-aware application: A machine learning based approach. International Journal of Smart Home, 9(1), 217-226.
[23] L. Zomlot, S. Chandran, D. Caragea, and X. Ou. Aiding intrusion analysis using machine learning. In Machine Learning and Applications (ICMLA), 2013 12th International Conference on (Vol.2, pp. 40-47). IEEE, 2013, December.
[24] Yadav, S. A., & Poongodi, T. (2021, April). A review of ml based fault detection algorithms in wsns. In 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM) (pp. 615-618). IEEE.
[25] Tweneboah-Koduah, S., Skouby, K. E., & Tadayoni, R. (2017). Cyber security threats to IoT applications and service domains. Wireless Personal Communications, 95, 169-185.
[26] Tewari, A., & Gupta, B. B. (2020). Security, privacy and trust of different layers in Internet-of-Things (IoTs) framework. Future generation computer systems, 108, 909-920.
[27] Srivastava, A., Gupta, S., Quamara, M., Chaudhary, P., & Aski, V. J. (2020). Future IoT-enabled threats and vulnerabilities: State of the art, challenges, and future prospects. International Journal of Communication Systems, 33(12), e4443.
[28] Itoo, F., Meenakshi, & Singh, S. (2021). Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection. International Journal of Information Technology, 13, 1503-1511.

*Edited by:* Venkatesan C
*Special issue on:* Next Generation Pervasive Reconfigurable Computing for High Performance Real Time Applications
*Received:* Sep 26, 2023
*Accepted:* Dec 5, 2023