WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Systematic Literature Review: Machine Learning Approaches for Enhancing IoT Security

Kontagora, Muhammad Mamman [a], Adeshina, Steve A.[b], Habiba, Musa [c]

[a] *PhD Candidate, Centre for Cyberspace Studies, Nassarawa State University, Keffi, Nassarawa State, Nigeria*
[b] *Professor, Department of Computer Engineering, Nile University, Abuja, Nigeria.*
[C] *Associate Professor, Department of Public and International Law, Nassarawa State University, Keffi, Nassarawa State, Nigeria.*

A B S T R A C T

Within the recent past, IoT devices, specifically those used in various industries, have greatly improved functionality and productivity. But, this growth process has also added a number of security threats, challenges, or risks. Conventional security approaches that require deep architectural integration on a rigid architecture do not sufficiently safeguard IoT systems since they are inherently diverse and complex. This systematic review aims at reviewing the possible use of machine learning (ML) approaches to mitigate these threats in IoT context. It gives an extended analysis of the specific susceptibilities of IoT devices and compares how a selection of techniques in the ML domain such as supervised, unsupervised, and reinforcement learning can be used to fight against threats. Among the benefits, the review specifies that ML can enable flexible data-driven security solutions in compliance with IoT needs. It is established that when data is analyzed in real-time and the model learns from its previous outcomes, it enhances the threat identification and prediction of threats by a huge margin. But it is a concern that there are still some major open issues that need to be addressed, including designing the ML model which makes it suitable for real-time threat detection, the incorporation of the explainable AI solution, and how to implement these models into resource-limited IoT devices. New directions in securing IoT as discussed in the review include the use of federated learning approaches to enhance privacy protection and the research into quantum-safe cryptographic techniques to handle future threats. Consequently, this review emphasizes the importance of continued research and development efforts to optimize ML applications for strengthening IoT defenses, providing constructive guidance for scholars, practitioners, and policymaking agencies who seek to improve IoT protection.

Keywords: Internet of Things (IoT), machine learning, threat analysis, security, prediction, explainable AI, federated learning.

1. Introduction

Many industries, including healthcare, manufacturing, housing, and transportation, are now manageable and controllable with accurate data decisions thanks to Internet of Things devices. Several global estimates of connected IoT devices have been made in the recent past, and based on them, the number of connected IoT devices is set to hit 30.9 billion in the year 2025 further, implying that individuals are using and incorporating these advancements at a very fast pace [1]. However, this extensive use also poses adverse security threats, given that IoT devices typically have few processing capabilities, memory, and battery capacity for employing standard security measures [2].

Security of IoT based systems is essential with the surge in the cases of cybercrimes that affect IoT systems. It also emerged that these devices may act as a gateway for attackers to access networks, mess with data and services [3]. Cybersecurity threats, including the Mirai Botnet attack, which targeted weak IoT systems and used these points of access to launch massive Distributed Denial of Service (DDoS) attacks, illustrate the potential hazards of insecurity in IoT systems [4]. Therefore, it is essential to fortify security protocols pertaining to devices presumed to be a part of the Internet of Things in order to avert the possible violation of individual rights and maintain the network's integrity [5], [6].

Therefore, ML has become one of the most promising approaches in addressing IoT security challenges since it involves the development of systems with self-learning capabilities [7]. One can illustrate the advantages of the ML approaches contrasting them with conventional rule-based security models: Compared to classical rule-based security mechanisms, ML technologies are capable of processing a large amount of data, and to pinpoint previously unknown security indicators of threats [8]. Machine learning is a set of techniques that is being widely implemented to tackle different forms of threats in IoT systems; supervised machine learning, unsupervised machine learning, and reinforcement learning [9, 10].

1.1 Objectives of the Review

The aim of this systematic literature review is to presents the current situation and state of the art of the threats, analysis, and prediction on the Internet of Thing (IoT) devices based on ML algorithms. The specific objectives are:

I. To identify and categorize several categories of threats that could be target to IoT devices and systems.

II. In this paper the authors provide a review of the used ML techniques for threat detection and prediction in IoT settings.

III. To evaluate the effectiveness of the applied models and the types of datasets for evaluating the model, as well as the most frequently used metrics.

IV. To review the limitations that has been presented to employ and evaluate the performance of ML in the contexts of both IoT security threats and to proposed new areas of research.

1.2 Research Questions

The review addresses the following research questions:

1) What are the current security threats to Internet of things (IoT) devices?
2) Generally, how well do these ML techniques work and what are measures, indices used to assess them?
3) What are the new trends and directions in this particular area?

2. Classification of Threats

IoT devices differ from one another due to their inherent connectivity capabilities, which expose them to various security risks in addition to their adaptability. They may be divided into three categories: threats from software, networks, and physical attack.

2.1 Physical Attacks

Some attacks are designed to take advantage of the physical attributes of IoT devices. Common forms include:

1) Tampering: interference with or destruction of the hardware fabrics or modules of the device with the intention to skew how it operates.
    a. Attackers may manipulate computer components to interfere with its regular operation or even hack into it to obtain privileged data [15].
    b. Side-Channel Attacks: Stealing the data by utilizing other signals like the amount of power consumed by a device or the electromagnetic emissions released by it [16].

2.2 Network Attacks

Network threats focus on the connections within the IoT devices. Key network threats include:

1. Man-in-the-Middle (MitM) Attacks: Intruders disrupt legitimate interactions between IoT devices and introduce unprovoked input data or control signals [17].
2. Distributed Denial of Service (DDoS) Attacks: Malicious IoT devices are employed to DDOS a target, thereby making it inaccessible. The Mirai botnet is one such instance, where the botnet targets IoT devices to launch big DDoS attacks [4], [18].
3. Eavesdropping: Malware intercepting normal data transmissions between the IoT devices that can cause data leakage [19].

2.3 Software Attacks

It acts in a similar fashion to malware, as it targets the flaws in the firmware of the device or particular application software. These include:

- Malware: A type of bot malware which can compromise IoT devices which can be leveraged by hackers to gain control of the device [20].
- Exploits and Vulnerabilities: It is used by the invaders to penetrate their target application by using the weaknesses to their advantage to cause some sort of disruption. Their proclivity for such vulnerability can be associated with poorly updated firmware [21].

2.4 Limitations to IoT Devices

IoT security is inherently challenging due to several factors:

2.4.1 Resource Constraints

A problem inherent in many IoT devices stems from the fact that the smallest nodes, or 'things,' in the Internet of Things tend to possess a minimal amount of processing power, memory and power, and are therefore unable to provide adequate security that is more conducive to more powerful environments [22].

2.4.2 Heterogeneity

The IoT system is made up of a number of heterogeneous objects that may have different functions and different operating systems and may use different types of protocols for the exchange of information. The last factor stems from the fact that the OSI model is a heterogeneous three-layered model that cannot have a uniform solution to a security problem [23].

2.4.3 Scalability

The need for security also diametrically increases as the number of connected nodes rises within the Internet of Things paradigm. Due to the fact that IoT networks are expected to consist of tens of billions of connected devices, the security solutions introduced therefore need to be highly scalable solutions that are capable of handling multiple devices and large volumes of data [24].

2.4.4 Privacy Concerns

By using IT services, IoT devices may contain personal information of the individuals. The most challenging aspect when it comes to the optimization of mobile applications is the balancing of data privacy and the provision of functionality. This is because; unauthorized collection or even hacks to the data on the server can result to severe violation of the privacy of an individual [25].

2.4.5 Lifecycle Management

The lifespan of an IoT device might be much longer than that of a typical computing device, yet they may not be as frequently updated with security patches. However, lifecycle management of security of such devices right from acquisition, usage, update, and disposal is another important aspect but commonly neglected [26].

2.5 Emerging Threats and Future Risks

The evolving IoT landscape brings new threats and risks:

2.5.1 AI-Powered Attacks

Artificial Intelligence (AI) can be used to enhance traditional attack methods, making them more adaptive and difficult to detect. AI can automate the discovery of vulnerabilities or optimize the coordination of large-scale attacks [27]. The problem of bias is one of the main ethical issues with AI and ML. When AI systems use historical data to learn, they may reinforce or even worsen preexisting biases in the data if such biases are present. Predictive police algorithms that unjustly target particular neighborhoods or automated employment systems that discriminate against particular demographics are just two examples of how these biases might appear. In order to increase the fairness of AI systems, this section will examine the nature of bias in AI, its effects, and mitigation techniques for bias.

2.5.2 Autonomous IoT Devices

With growing independence due to increasing integration of IoT devices, the prospects of its malicious manipulation only materialize. The malicious attacks that target the autonomy of these IoT systems can cause a range of security and safety issues; the systems can be tricked into performing specific actions that were not intended [28].

2.5.3 Quantum Computing

This is a contentious issue given that potential future development of quantum computing may crack current cryptographic techniques hence debilitating IoT security. Addressing the problem of crafting a post-quantum cryptographic approach towards IoT is imperative to build resilient IoT systems [28].

3. An Overview of Machine Learning Techniques

Threat analysis and prediction in the IoT context is made easier and more effective by the tools provided by Machine Learning (ML). The learning can be divided into supervised, unsupervised, semi-supervised or reinforcement learning techniques. In its turn, each category has its own advantages when the process of security threats detection and prevention is concerned.

3.1 Supervised Learning

Supervised learning requires a model to be learned where the results are already categorized or classified. This approach is highly effective in identifying known threats:

1) Classification: Applied to sort threats under certain predetermined type. Based on the nature of the studied problem, common methods include Support Vector Machines (SVMs), Decision Trees, and Neural Networks [30], [31]. For instance, SVMs has been used in detecting misbehavior within IoT networks [32].
2) Regression: Used when the dependent variable is a metric one, i. e. , a degree of the attack's effect or probability. Linear regression models can make predictions of the intensity that possible threats pose in order to formulate accurate estimations from past incidences [33].

3.2 Unsupervised Learning

Unsupervised learning deals with unlabeled data, making it suitable for discovering unknown threats:

- Clustering: Clustering like data in a particular category as they are closer to each other. Machine learning algorithms like k-Means, DBSCAN and Hierarchical Clustering can point to a case of a new threat on the horizon through their analysis of abnormal patterns [34]. To note, clustering has been applied to identify the traffic abnormalities in IoT networks [35].
- Anomaly Detection: Detects cases and finds out if there are abnormalities. For example, to detect the anomalies in IoT data streams Principal Component Analysis (PCA) and Isolation Forests are usually applied [36].

3.3 Semi-Supervised Learning

Semi-supervised learning combines labeled and unlabeled data, leveraging small amounts of labeled data to improve learning accuracy:

- Self-Training: Incorporates initial labeled data to sort other data which are unsorted, with a constant feedback loop to enhance the conclusion [37].
- Graph-Based Methods: Uses graphs to capture relative data values which enable identification of botnets or malware trespassing across constrained IoT devices [38].

3.4 Reinforcement Learning

Reinforcement learning involves training models through interactions with the environment, learning optimal actions to maximize cumulative reward:

1) Q-Learning: A value function of learning policy in environment stochastic and non-stationary: A method without models [39]. For modifying security in IoT networks in accordance with emergent threats, Q-Learning has been incorporated [40].
2) Deep Reinforcement Learning: Proposes a proactive preventive threat detection system as a combination of neural network with reinforcement learning principles [41].

3.5 ML applications for Internet of Things security

Image and pattern recognition, classification, anomaly detection, and prevention approaches are used in IoT applications to improve security based on the ML technique.

3.5.1 Intrusion Detection Systems (IDS)

Intrusion Detection Systems (IDS) use ML models to identify unauthorized access or anomalies in network traffic:

i. Signature-Based IDS: Analyzes the data and tries to match the traffic patterns against a database that contains well-known threat signatures. Despite a high level of efficiency when it comes to the known attacks, it is less capable of dealing with new ones [42].
ii. Anomaly-Based IDS: Is able to monitor for irregularities from normal traffic patterns that would help in the identification of new or advanced threats [43]. In the interests of IoT network anomaly-based IDS, Random Forest and Autoencoders have been applied [44].

3.5.2 Malware Detection

ML models can analyze the behavior and characteristics of software to detect malware:

- Static Analysis: Does not involve the running of the code, but rather employs methods that analyze code characteristics. Algorithms like Decision Trees, and Logistic Regression are able to categorize the code into either malicious or benign basing the decision on the features of the code [45].
- Dynamic Analysis: The ability to control and observe activities of the software during their work and search for a violation of set rules. The traditional features are analysed by using sophisticated Deep Learning models like Convolutional Neural Networks (CNNs) for malware detection from dynamic features [46].

3.5.3 Botnet Detection

ML techniques are effective in detecting and mitigating botnets that leverage IoT devices for coordinated attacks:

1. Flow-Based Detection: Scans the traffic to detect any signs of botnet being at work. With the help of techniques such as Clustering and SVMs one could classify traffic patterns relevant to botnets [47].
2. Behavior-Based Detection: Detects abberant behavior in the devices which may signify the presence of a botnet. The implementation of Long Short-Term Memory (LSTM) networks will enable capturing the temporal dynamics of behavior to set alarms for botnet actions [48].

4. Performance Evaluation Metrics

Evaluating the performance of ML models in IoT security involves several metrics to ensure accuracy and reliability:

4.1 Accuracy and Precision

- Accuracy: An index of the degree of threat accuracy, calculated as the ratio of valid threat identification to total threat identification. The above results show high accuracy meaning that the model is well capable of discriminant threats from non-threats [49].
- Precision: Assessment of the ratio of true positive and reflecting the model's capacity to filter out False Positives [50].

4.2 Recall and F1-Score

1. Recall: Measures the consistent detection of genuine threats, which shows how many mentions of threats were actual realistic threats the model successfully identified[33].
2. F1-Score: The average between both precision and recall, ensuring that any model yielded is balanced [52].

4.3 Receiver Operating Characteristic (ROC) and the Area under the Curve (AUC).

1. ROC Curve: Details the true positive rate against the false positive rate intended to demonstrate the trade-off between sensitivity and specificity [53].
2. AUC: Reflects measurement for the proportion of actual positives; it reflects the overall performance of the model [54].

4.4 Challenges and Limitations of Machine Learning Algorithm Highlighted in IoT Security

Despite the advantages, there are several challenges in applying ML to IoT security:

Table 4.1: Challenges and Limitations of ML in IoT

Quantity and Quality of Data	The IoT is characterized by extensive heterogeneity in terms of the number of connected objects and the type and quantity of data it produces. An important note about datasets is the need to maintain high quality and representativeness in them, as this is the foundation for the effectiveness of working with ML models.
Computational Constraints	Restricted computational capabilities limit the IoT devices to use mainstream shallow machine learning and also require efficient models.
Adversarial Attacks	The adversarial attacks are known to happen in ML models while suggesting that the given inputs are manipulated to make the incorrect predictions by the model.
Generalization and Adaptability	While training the model with some datasets may provide good solutions to be implemented on the specific environment or in terms of threat detection, the same models might not work in different environments or under new threats and hence the ML models requires a continuous learning process.
Privacy Concerns	Despite these benefits, the integration of ML with IoT poses privacy issues, since the ML models can perform data processing on the private information of individuals. Preservation of data, ownership, and clients' confidentiality as well as respect for legislation and basic norms are crucial.

Source: [55],[56],[57],[58]

5. Performance Metrics and Evaluation

When machine learning is taken into account in relation to anomaly detection in the Internet of Things (IoT) sector, the evaluation starts to make sense. The metrics offer methods for assessing the effectiveness of tried-and-true anomaly detection strategies, the top-performing algorithms, and the suitability of such algorithms. A synopsis of the most often used performance metrics is provided below: Important performance analysis methodologies and the function of each metric are also covered in this section.

5.1 Overview of Metrics used to Evaluation Machine Learning Models

5.1.1 Accuracy

Accuracy is the purest of the metrics, the simplest method that shows the ratio of the correctly identified both normal and anomalous instances to the total number of instances. It displays a comparison metrics between two or more groups or a unified measure of group performance. Potentially proactive in datasets that are unbalanced and have a comparatively higher proportion of normal cases than anomalous ones [33].

It is given by:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (1)$$

Where; *TP* = True Positives, *TN* = True Negatives, *FP* = False Positives and *FN* = False Negatives

5.1.2 Precision and Recall

Precision measures the proportion of correctly identified anomalies to the total number of instances classified as anomalies. In situations when it is necessary to compromise the amount of false positives and false negatives, these measures are relatively extremely helpful. While recall aims to identify every true anomaly, precision allows one to determine the extent to which detected anomalies are significant. Even yet, it is crucial to take into account both in order to see the big picture because improper usage of one might have an impact on the other:

$$Precision = \frac{TP}{TP+FP} \qquad (2)$$

Recall (or Sensitivity) measures the proportion of actual anomalies that are correctly identified:

$$Recall = \frac{TP}{TP+FN} \qquad (3)$$

5.1.3 F1-Score

The F1-Score is the harmonic mean of Precision and Recall and is used to balance them. This simplifies the process of comparing several models by combining Precision and Recall into a single figure. In the event of severely skewed datasets, the precision and recall trade-off area might not be correctly exposed [35].:

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \qquad (4)$$

5.1.4 Area Under the Receiver Operating Characteristic Curve (AUC-ROC)

Plotting the true positive rate against the false positive rate at different probability thresholds yields the AUC-ROC, a graphical depiction of a binary classifier system's diagnostic capacity. It plots the True Positive Rate (TPR) against the False Positive Rate (FPR) at various threshold settings:

$$TPR = \frac{TP}{TP + FN}$$

$$FPR = \frac{FP}{FP + TN}$$

The best model performance is indicated by an AUC that is closer to 1, on a scale or range of 0 to 1. It provides a general indicator of how well the model can categorize patterns into different classes. It does not utilize genuine market prices since doing so would incur costs related to misclassification[36].

5.2 Typical Experimental Set-Up and Validation Techniques

5.2.1 Cross-Validation

In order to assess a model's efficacy using multiple splits of the observed dataset, cross-validation entails randomly dividing the observed dataset into the training and testing datasets. K-fold A common technique called cross-validation involves splitting each dataset into K subsets and training and validating the machine learning algorithm K times. During each iteration, one of the K subsets is used as a test set. It lessens the likelihood of getting excellent training data scores that do not convert to better results on other data. In addition to being computationally and statistically demanding, it can also take a long time with huge datasets [37].

5.2.2 Hold-Out Validation

Two distinct sets of the dataset are used for Hold-Out Validation: Typically, there are two types of datasets: the training set and the testing set. The proportion of the data that is split into 70–80% for training and 20–30% for testing must be equal. Compared to cross-validation, it takes less computing resources and is simpler to deal with or implement. The split may have an impact on these measurements, and in certain cases, the findings may be more or less accurate [38].

6. Future Research Directions

IoT security as a field is developing and there is still a lot of opportunities to develop new approaches that will improve the effectiveness of threat analysis and forecasting using machine learning. To sum up, some critical directions of further studies are described in this section, stressing the need for the development of new approaches to solve issues and threats, which appeared in the world of IoT security and to use new opportunities effectively.

6.1 Emerging Strategies for Detecting Advanced Threats

6.1.1 Antibodies used in Federating Learning for IoT security

FL is a collective learning technique in which model is learned at multiple devices without sharing the actual data. The simplicity of this technique affords large privacy benefits since the data remains highly localized and can be used in any context where data privacy and bandwidth are issues as is the case in IoT. Future research can explore:

- Scalable FL Frameworks: Establishing the federated learning frameworks specific to the different types of IoT networks with needed solutions before and after the model aggregation, focusing on the communication message integrity and security from adversarial threats [60].

- Secure Aggregation Protocols: Creating a way for an aggregation algorithm to maintain the privacy and accuracy in the aggregation of models from local models to the global model.

6.1.2 Explainable Artificial Intelligence in Threat Identification

In general, Explainable AI or XAI focuses on how to enable the ML model to be explainable, that is, to explain what the model has learned and why it made a particular decision. This can enhance trust and transparency in threat detection systems:

- Interpretable Models: The formulation of accurate models that will give satisfactory explanations for the predicted threats will enable the security analyst to independently verify and thus have confidence in the identified threats.

- Post-Hoc Explanations: Applying post Hoc explanation strategies to intricate large models where they use the feature importance extraction method and Local Interpretable Model-Agnostic Explanations (LIME).

6.1.3 Adversarial Machine Learning

Adversarial Machine Learning focuses on making ML models robust against adversarial attacks where inputs are intentionally manipulated to deceive the model:

- Adversarial Training: Using adversarial examples during training to make the model better prepared for such edits or insertions.

- Detection and Mitigation: Creating techniques for identifying and combating adversarial attacks in real time so as to achieve precise threat indication [62].

6.2 Integration of Emerging Technologies

6.2.1 A blockchain for the security of IoT

Blockchain technology provides a decentralized and immutable ledger, offering potential solutions for enhancing IoT security:

- Secure Data Sharing: Implementing blockchain to establish secured and immutable channels for conveying data between IoT devices and maintaining the data's sanctity and inviolability, as well as restraining unauthorized access to it.

- Smart Contracts: Security policies to be written in smart contracts and used to automate the provisioning of security across IoT networks, thus minimizing the occurrence of human mistakes [63].

6.2.2 Quantum-Resistant Cryptography

In the progression of quantum computing, current forms of cryptography for potential vulnerabilities have been threatened. Research into quantum-resistant cryptography is essential for future-proofing IoT security:

- Post-Quantum Algorithms: This would entail creating new cryptographic algorithms that would be hard to break by quantum attacks while at the same time ensuring that IoT devices will be safe in future from possible attacks.

- Hybrid Cryptographic Solutions: Developing modern cryptosystems capable of protecting users' information in both present and future with quantum-safe options as a progressive approach [64].

6.3 Enhanced Privacy-Preserving Techniques

6.3.1 Differential Privacy

Differential privacy techniques aim to provide privacy guarantees while allowing the extraction of useful insights from data:

- Privacy Mechanisms: Introducing privacy mechanisms that can add some form of noise into data or queries, maintain privacy while improving the threat identification success rate.

- Evaluation Metrics: Indeed, one of the promising yet major challenges is to devise satisfactory benchmarks that could be used to assess privacy-loss trade-offs in differential privacy deployments for IoT data [65].

6.3.2 Homomorphic Encryption

Homomorphic encryption allows computations on encrypted data without decrypting it, preserving privacy throughout the data processing pipeline:

- Efficient Algorithms: It is also directly related to the current challenges of coming up with effective homomorphic encryption algorithms which are practical to use in resource-limited IoT settings.

- Scalable Solutions: Extending homomorphic cryptographic protocols for application on large IoT networks with emphasis on confidentiality/computation trade-off [66].

6.4 Context-Aware and Adaptive Security Systems

6.4.1 Context-Aware Threat Detection

Context-aware systems consider environmental and situational factors to improve threat detection accuracy:

- Dynamic Context Analysis: To utilize the proposed approaches and methods to introduce methods for aware detection of the context in which IoT devices are used and of adjusting the corresponding detection mechanisms according to the identified context.

- Adaptive Models: There are new categories of models that are adaptive and can adjust their functions based on changes such as the network standards and users' actions [67].

6.4.2 Adaptive Authentication Mechanisms

Adaptive authentication adjusts security requirements based on the risk profile of the current situation:

- Risk-Based Authentication: Adapting to the use of risk Based Authentication Protocols that enhances the security of the system on instances of malicious activity.
- Multi-Factor Authentication: Include reversible and easily manageable multi-factor solutions that use machine learning to implement reliable and context-based approaches [68].

6.5 Cross-Domain Threat Intelligence

6.5.1 Collaborative Threat Intelligence

Collaborative threat intelligence involves sharing threat information across different organizations or sectors to improve detection capabilities:

- Threat Sharing Frameworks: Furthering the identification of reference models for safe and optimal information exchange regarding threats within the context of achieving synergy with other invested parties.
- Automated Sharing Protocols: Developing 'use case templates' that allow for the establishment of real-time, automated threat intelligence sharing between organizations, leading to more timely and relevant information sharing [69].

6.5.2 Super Learning for Threat Identification

Transfer learning allows models trained in one domain to be applied to another, facilitating the detection of threats across different IoT environments:

- Domain Adaptation: Analyzing and comparing approaches of fine-tuning trained models to new, but related, IoT environments, useful in enhancing the generalization of threat detection systems.
- Cross-Domain Datasets: Selecting datasets common in different domains for improving transfer learning methods and increasing their accuracy in various applications [2].

6.6 Ethical and Societal Implications

6.6.1 Ethical Considerations in AI-Driven Security

AI-driven security solutions must address ethical considerations to ensure fair and responsible use:

3. Bias Mitigation: Making sure that no users are unfairly treated by the algorithms in use in the machine learning models through eliminating bias.
4. Transparency and Accountability: Giving consideration to the importance of transparency in the AI decision-making mechanisms and providing oversight and culpability regarding the AI-secured systems.

6.6.2 Societal Impact of IoT

The widespread adoption of IoT devices has significant societal implications, particularly regarding security:

- Privacy Rights: Studying measures that strengthen security alongside introducing measures that protect people rights not to be limited by IoT security policies.
- Public Awareness: It enhances people's awareness about IoT security threats and encouraging people to apply the correct methods to use smart devices securely [72].

7. Conclusion

In the new layouts of Smart world, IoT products have incorporated in human and daily works. This has positive impacts on the work efficiency and other operation aspects. However, this advent brings in a large number of threats to the security aspect which in turn require effective and dynamic security systems. However, with the onset of ML, the threat analysis and prediction of these threats has become much more sophisticated and not merely restricted to conventional security tools.

The crux of security issues surrounding IoT is perhaps best solved by the use of machine learning as a novel means of addressing the issue. The core idea, therefore, is to use the advantages and opportunities of the ML approach to create higher adaptability, performance, and resistance to threats for security systems, preserving the pace of development of IoT security threats. This systematic review suggests more emphasis placed on the development and application of algorithmic research and policy measures in the prevention of threats posed by IoT vulnerabilities through the effective use of ML.

References

1. Statista. (2021). Number of IoT connected devices worldwide 2019-2030. [Online]. Available: https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/
2. Juniper Research. (2022). The Internet of Things: How IoT is Revolutionising Industry. [Online]. Available: https://www.juniperresearch.com/researchstore/connected-devices/iot-market-research-report
3. S. Sicari, A. Rizzardi, L. A. Grieco, and A. Coen-Porisini, "Security, privacy and trust in Internet of Things: The road ahead," Computer Networks, vol. 76, pp. 146-164, 2015.
4. E. Bertino and N. Islam, "Botnets and Internet of Things Security," Computer, vol. 50, no. 2, pp. 76-79, 2017.
5. R. H. Weber, "Internet of Things – New security and privacy challenges," Computer Law & Security Review, vol. 26, no. 1, pp. 23-30, 2010.
6. Krebs on Security. (2016). DDoS on Dyn Impacts Twitter, Spotify, Reddit. [Online]. Available: https://krebsonsecurity.com/2016/10/ddos-on-dyn-impacts-twitter-spotify-reddit/
7. H. Suo, J. Wan, C. Zou, and J. Liu, "Security in the Internet of Things: A review," in 2012 International Conference on Computer Science and Electronics Engineering, 2012, vol. 3, pp. 648-651.
8. J. H. Ziegeldorf, O. G. Morchon, and K. Wehrle, "Privacy in the Internet of Things: Threats and challenges," Security and Communication Networks, vol. 7, no. 12, pp. 2728-2742, 2014.
9. N. Moustafa, J. Hu, and J. Slay, "A holistic review of Network Anomaly Detection Systems: A comprehensive survey," Journal of Network and Computer Applications, vol. 128, pp. 33-55, 2019.
10. L. Chen, L. Lin, Q. Xu, Y. Zhou, and M. Xian, "An overview of machine learning in Internet of Things," Concurrency and Computation: Practice and Experience, vol. 32, no. 15, e5287, 2020.
11. Y. Meidan et al., "Detection of unauthorized IoT devices using machine learning techniques," Procedia Computer Science, vol. 135, pp. 265-272, 2018.
12. M. Ammar, G. Russello, and B. Crispo, "Internet of Things: A survey on the security of IoT frameworks," Journal of Information Security and Applications, vol. 38, pp. 8-27, 2018.
13. P. Kumar and P. K. Srivastava, "Machine Learning in the Internet of Things: A Systematic Review," in 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), 2019, pp. 9-14.
14. A. A. Malik, M. S. Munir, and W. H. Mahmood, "An analysis of security challenges in Internet of Things (IoT)," in 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), 2019, pp. 1-5.
15. M. A. Alrawashdeh and C. Purdy, "Cyber Security Attacks on Smart Grids," IEEE Transactions on Smart Grid, vol. 10, no. 1, pp. 114-125, 2019.
16. M. D. Joye, M. Tunstall, and M. A. Hasan, "Practical Side-Channel Attacks against Cryptographic Software," in International Conference on Smart Card Research and Advanced Applications, 2012, pp. 1-13.
17. Y. Yang, L. Wu, G. Yin, L. Li, and H. Zhao, "A Survey on Security and Privacy Issues in Internet-of-Things," IEEE Internet of Things Journal, vol. 4, no. 5, pp. 1250-1258, 2017.
18. G. Blanc, Z. Zhang, P. Naylor, A. Finamore, and F. Rousseau, "Botnet Command and Control Traffic Detection through Network Flow Clustering," Journal of Network and Computer Applications, vol. 50, pp. 46-56, 2015.
19. Z. Jin, H. Liu, and Y. Zhang, "A Study of Eavesdropping in IoT," in 2016 IEEE International Conference on Internet of Things (iThings), 2016, pp. 221-226.
20. E. Fernandes, J. Jung, and A. Prakash, "Security Analysis of Emerging Smart Home Applications," in 2016 IEEE Symposium on Security and Privacy (SP), 2016, pp. 636-654.
21. L. Davi, A. Dmitrienko, A.-R. Sadeghi, and M. Winandy, "Privilege Escalation Attacks on Android," in 13th International Conference on Information Security, 2010, pp. 346-360.
22. M. Ammar, G. Russello, and B. Crispo, "Internet of Things: A survey on the security of IoT frameworks," Journal of Information Security and Applications, vol. 38, pp. 8-27, 2018.

23. S. Neshenko, E. Bou-Harb, J. Crichigno, G. Kaddoum, and N. Ghani, "Demystifying IoT Security: An Exhaustive Survey on IoT Vulnerabilities and a First Empirical Look on Internet-Scale IoT Exploitations," IEEE Communications Surveys & Tutorials, vol. 21, no. 3, pp. 2702-2733, 2019.
24. M. Ammar, G. Russello, and B. Crispo, "Internet of Things: A survey on the security of IoT frameworks," Journal of Information Security and Applications, vol. 38, pp. 8-27, 2018.
25. R. Roman, J. Lopez, and M. Mambo, "Mobile Edge Computing, Fog et al.: A Survey and Analysis of Security Threats and Challenges," Future Generation Computer Systems, vol. 78, pp. 680-698, 2018.
26. K. D. Singh and M. Khari, "Review on Security Challenges in Internet of Things (IoT)," in 2019 2nd International Conference on Power Energy, Environment and Intelligent Control (PEEIC), 2019, pp. 291-297.
27. F. Mohammadi, S. Zeadally, and L. Anjum, "AI-Powered Cyber Threat Intelligence for Secure IoT Ecosystems," IEEE Internet of Things Magazine, vol. 2, no. 2, pp. 10-15, 2019.
28. B. Xu, L. Da Xu, H. Cai, C. Xie, J. Hu, and F. Bu, "Ubiquitous Data Accessing Method in IoT-Based Information System for Emergency Medical Services," IEEE Transactions on Industrial Informatics, vol. 10, no. 2, pp. 1578-1586, 2014.
29. M. Mosca, "Cybersecurity in an Era with Quantum Computers: Will We Be Ready?," IEEE Security & Privacy, vol. 16, no. 5, pp. 38-41, 2018.
30. T. V. L. N. Mallikarjuna Reddy, B. Krishna Reddy, A. R. Urs, "IoT Threat Detection with Supervised Learning," IEEE Access, vol. 8, pp. 153742-153751, 2020.
31. A. Koubaa, M. Nazir, S. Cheikhrouhou, A. Ammar, and E. Tovar, "A Comprehensive Survey on Machine Learning for Cyber-Physical System Security," Sensors, vol. 21, no. 8, pp. 1-31, 2021.
32. D. U. Reddy and K. V. Prasad, "Intrusion Detection in IoT Networks using Support Vector Machines," in 2021 International Conference on Advances in Computing, Communication, and Control (ICAC3), 2021, pp. 1-6.
33. G. Omar, A. B. Hassan, and K. A. Bin-Khafan, "Predictive Maintenance of IoT Devices using Regression Analysis," Journal of Engineering Research and Applications, vol. 12, no. 5, pp. 45-50, 2022.
34. D. Mahmood and S. A. R. Z. Bokhari, "Unsupervised Learning Approaches for Threat Detection in IoT," in 2020 IEEE International Conference on Data Science and Data Intensive Systems (DSDIS), 2020, pp. 110-116.
35. J. Shin and J. Park, "IoT Security Threat Detection Based on Network Traffic Clustering," Journal of Communications and Networks, vol. 22, no. 1, pp. 45-53, 2020.
36. P. Singh, A. Jain, and A. K. Jaiswal, "Anomaly Detection in IoT Using Principal Component Analysis," in 2021 International Conference on Innovations in Information Technology (IIT), 2021, pp. 1-6.
37. J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. Waltham, MA: Morgan Kaufmann, 2011.
38. W. Zhou, Y. Qian, M. Keck, and R. Lu, "Graph-Based Semi-Supervised Learning for IoT Security," IEEE Internet of Things Journal, vol. 8, no. 3, pp. 2037-2048, 2021.
39. R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed. Cambridge, MA: MIT Press, 2018.
40. Y. Pan and X. Li, "A Deep Reinforcement Learning Framework for Security in IoT," IEEE Access, vol. 8, pp. 1235-1248, 2020.
41. Y. Liu et al., "Deep Reinforcement Learning for IoT Security," in 2020 IEEE International Conference on Network Protocols (ICNP), 2020, pp. 1-11.
42. G. Vasudevan and S. Anil, "Machine Learning-based Signature Detection for IoT Devices," in 2021 IEEE International Conference on Machine Learning and Applications (ICMLA), 2021, pp. 68-74.
43. R. Mitchell and I.-R. Chen, "A Survey of Intrusion Detection Techniques for Cyber-Physical Systems," ACM Computing Surveys, vol. 46, no. 4, pp. 1-29, 2014.
44. H. Wang, H. Wang, and W. Wang, "IoT Security Anomaly Detection Using Autoencoders," in 2021 IEEE International Conference on Consumer Electronics (ICCE), 2021, pp. 1-6.
45. W. S. McCulloh and W. Pitts, "A Logical Calculus of the Ideas Immanent in Nervous Activity," Bulletin of Mathematical Biophysics, vol. 5, no. 4, pp. 115-133, 1943.
46. M. Z. Ahmad et al., "Dynamic Analysis for IoT Malware Detection Using Deep Learning," Future Generation Computer Systems, vol. 124, pp. 245-258, 2021.
47. Z. A. Baig et al., "Botnet-Based Distributed Denial of Service (DDoS) Attacks on Web Servers: Classification and Art," Journal of Network and Computer Applications, vol. 36, no. 1, pp. 1-14, 2013.

48. P. Kolios, V. Vassiliou, and V. Sassaris, "Botnet Detection Using Behavior-Based Models and LSTM Networks," IEEE Transactions on Network and Service Management, vol. 17, no. 3, pp. 1204-1217, 2020.
49. A. Mishra and R. Jain, "Accuracy Measures for Machine Learning: An Overview," in 2021 IEEE International Conference on Data Science and Data Intensive Systems (DSDIS), 2021, pp. 73-79.
50. F. Provost and T. Fawcett, "Robust Classification for Imprecise Environments," Machine Learning, vol. 42, no. 3, pp. 203-231, 2001.
51. J. Davis and M. Goadrich, "The Relationship Between Precision-Recall and ROC Curves," in Proceedings of the 23rd International Conference on Machine Learning (ICML), 2006, pp. 233-240.
52. P. G. L. M. Ferris, "The F1-Score: An Incomplete Evaluation Metric for Binary Classifiers," Data Mining and Knowledge Discovery, vol. 8, no. 2, pp. 189-202, 2004.
53. T. Fawcett, "An Introduction to ROC Analysis," Pattern Recognition Letters, vol. 27, no. 8, pp. 861-874, 2006.
54. C. D. Manning, P. Raghavan, and H. Schütze, Introduction to Information Retrieval, Cambridge, UK: Cambridge University Press, 2008.
55. G. Farinella, G. Di Bello, A. Furnari, and S. Battiato, "A Review of Challenges and Opportunities in Processing and Mining Large-scale IoT Data," IEEE Access, vol. 9, pp. 160823-160847, 2021.
56. H. Suo, J. Wan, C. Zou, and J. Liu, "Security in the Internet of Things: A Review," in 2012 International Conference on Computer Science and Electronics Engineering, 2012, pp. 648-651.
57. I. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and Harnessing Adversarial Examples," in Proceedings of the 2015 International Conference on Learning Representations (ICLR), 2015, pp. 1-12.
58. M. Fredrikson et al., "Privacy in the Age of Machine Learning," Communications of the ACM, vol. 60, no. 10, pp. 58-66, 2017.
59. J. K. L. Xie and H. Ling, "Generalization in Machine Learning: A Survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 41, no. 9, pp. 1-14, 2022.
60. Y. Lu, X. Huang, Y. Dai, S. Maharjan, and Y. Zhang, "Blockchain and Federated Learning for Privacy-Preserved Data Sharing in Industrial IoT," IEEE Transactions on Industrial Informatics, vol. 16, no. 6, pp. 4177-4186, 2020.
61. M. T. Ribeiro, S. Singh, and C. Guestrin, "Why Should I Trust You? Explaining the Predictions of Any Classifier," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2016, pp. 1135-1144.
62. N. Papernot et al., "Practical Black-Box Attacks against Deep Learning Systems Using Adversarial Examples," in Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, 2017, pp. 506-519.
63. Y. Zhang, L. Wu, and G. Wei, "Blockchain-Based Systems and Applications: A Survey," IEEE Access, vol. 7, pp. 117134-117151, 2019.
64. P. W. Shor, "Algorithms for Quantum Computation: Discrete Logarithms and Factoring," in Proceedings of the 35th Annual Symposium on Foundations of Computer Science, 1994, pp. 124-134.
65. R. Bassily and A. Smith, "Local, Private, Efficient Protocols for Succinct Histograms," in Proceedings of the 47th Annual ACM Symposium on Theory of Computing (STOC), 2015, pp. 127-135.
66. P. Paillier, "Public-Key Cryptosystems Based on Composite Degree Residuosity Classes," in Advances in Cryptology (EUROCRYPT), 1999, pp. 223-238.
67. H. Zhang and Z. Liu, "A Context-Aware Security Framework for IoT Applications," IEEE Access, vol. 6, pp. 17613-17621, 2018.
68. A. J. Mendez, D. V. Sanchez, and S. Carrasco, "Risk-Based Adaptive Authentication for IoT," in Proceedings of the 2020 IEEE Conference on Communications and Network Security (CNS), 2020, pp. 1-8.
69. K. K. R. Choo, "The Cyber Threat Landscape: Challenges and Future Research Directions," Computers & Security, vol. 32, pp. 719-731, 2013.
70. S. J. Pan and Q. Yang, "A Survey on Transfer Learning," IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 10, pp. 1345-1359, 2010.
71. R. Danks and C. London, "Algorithmic Bias in Autonomous Systems," Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT), 2021, pp. 127-138.
72. C. Maple, "Security and Privacy in the Internet of Things," Journal of Cyber Policy, vol. 2, no. 2, pp. 155-184, 2017.