# Machine Learning Based Cyber Security Technique For Detection Of Upcoming Cyber Attacks

**Asma Nikhat**
Assistant professor
Department of Computer science
Government College (A), Kalaburagi

**Abstract**
Cyber attacks are prevalent in the age of the Internet. Each year, both the quantity and severity of cybercrimes increase. Protection against cyber-attacks has become a primary responsibility, Significant in the internet society of today. However, providing cyber security is a highly difficult task that requires experience in the field of attacks and the ability to evaluate the possibility of threats. The continual evolution of cyber attacks is the biggest challenge in this industry. In the increasingly digital world, the frequency and sophistication of cyber attacks continue to rise, posing significant threats to individuals, organizations, and nations. Traditional cyber security measures often fall short in identifying and mitigating novel and evolving threats. This paper presents a machine learning-based cyber security technique designed to detect and predict upcoming cyber attacks. The proposed system leverages advanced machine learning algorithms to analyze vast amounts of network data and identify patterns indicative of potential threats.

Experimental results demonstrate the system's capability to detect a wide range of cyber attacks, including zero-day exploits, phishing, and distributed denial-of-service (DDoS) attacks. The proposed technique shows promising results in terms of detection accuracy, false positive rate, and response time, highlighting its potential as a robust solution for proactive cyber defense. This research contributes to the field of cyber security by providing a scalable, adaptive, and efficient approach to threat detection. Future work will focus on enhancing the model's scalability, integrating it with existing security infrastructure, and conducting extensive field tests to validate its performance in real-world scenarios. The most effective method of protection from these attackers is to conduct vulnerability analysis against newly emerging threats for the systems we use and to rectify identified vulnerabilities. In this research paper, the weaknesses of wireless communication towards remote connection usage of the mini electric autonomous vehicle were investigated, which we developed and produced its mechanics, electronics, and software.

**Keywords:** Cyber security, Machine Learning, Artificial Intelligence (AI), Cyber Attacks, Supervised Learning Unsupervised Learning Reinforcement Learning Deep Learning

## I. Introduction

As more people depend on and use the Internet, companies, governments, and even banks have moved their activities online. Therefore, cyber defenses are weakened. Cyber attacks are defined as any intentional attempt to obtain unauthorized access to a target's computer network. Businesses, governments, and other financial institutions such as banks are regular targets of cyber attacks designed to steal valuable data or demand a ransom in return for access.

As technology rapidly advances, research on robots is intensifying, with the range of applications for robots expanding continuously. The widespread use of the Internet has started to impact our phones, vehicles, and even our homes. Robots and autonomous systems, which are beginning to adapt to human life, have also been integrated into land vehicles in recent years. These autonomous technologies adapted to land vehicles bring features that enhance the comfort of human life.

In this study, we embark on a pioneering journey to address the critical cyber security challenges faced by autonomous vehicle systems. By focusing on the prevalent vulnerabilities within wireless communication networks, we explore the real-world implications of various cyber attacks like Deauth, DoS, DDoS, and MitM. We implement and test the efficacy of advanced machine learning algorithms, particularly the random forest algorithm, in detecting these threats. The implications of our findings are far-reaching, offering vital insights and tools for stakeholders in the rapidly evolving landscape of autonomous driving technologies. Key components of the system include a comprehensive data collection framework that gathers real-time data from various network sources, a feature extraction module that processes and transforms raw data into meaningful features, and a machine learning model that detects anomalies and predicts potential attacks. The model is trained on a diverse dataset comprising historical attack data and benign network activity, enabling it to distinguish between normal and malicious behavior.

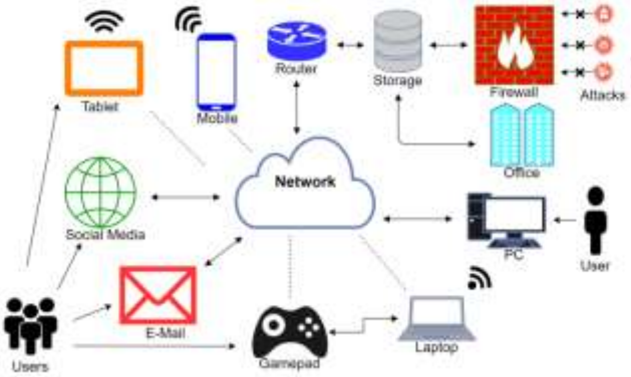
## Discussion

The utilization of autonomous vehicles within the scope of Industry 4.0 has become an increasingly popular area of interest. However, autonomous vehicles possess several factors that make them vulnerable to cyber attacks. Consequently, the issue of autonomous vehicle security has been seriously addressed by researchers in recent times. In this context, Artificial Intelligence-based attack detection systems have been developed. Expanding on the role of machine learning in cyber security, we delve into a detailed analysis of its applications, challenges, and future prospects. Machine learning (ML), a subset of artificial intelligence, has become a cornerstone in modern cyber security strategies. Its ability to process and learn from large volumes of data enables it to identify patterns and anomalies that would be impossible for humans to detect efficiently. The methodology begins with designing and constructing a mini electric autonomous vehicle tailored to simulate real-world autonomous systems. This involves detailed mechanical, electronic, and software configurations to ensure the vehicle accurately represents larger, more complex systems. The next phase simulates various cyber attacks, including Deauth, DoS, DDoS, and MitM attacks.

## Analysis

### Detection of attacks using artificial intelligence

AI Intelligence algorithms on network traffic. The recorded network traffic is passed through various AI algorithms. In this study, the Artificial Intelligence-based attacker detection model consists of four stages. In the first stage, the data obtained from network traffic is processed through data preprocessing steps to create a suitable dataset. This dataset is divided into 10-millisecond time intervals before loading onto the model to enhance the accuracy of the algorithms. In the second



## Discussion and results

In this study, the effectiveness of the Random Forest algorithm for intrusion detection was evaluated based on the analysis of network traffic data. The goal was to assess the ability of the algorithm to accurately classify packets as either attack or non-attack. The results obtained from the evaluation were discussed in relation to the system's performance in terms of accuracy, precision, recall, and false positive rate.

*Supervised machine learning in cyber security*

Supervised machine learning in cyber security is used to classify data or predict outcomes. It uses labeled datasets to train algorithms and define the variables to be assessed for correlations, with the input and outputs specified. As part of the cross-validation process, when input data is fed, the model adjusts its weights until it has been fitted appropriately to avoid over fitting or under fitting. Identifying unique labels of network risks, such as scanning and spoofing. Predicting or classifying a target variable for a specific security threat (e.g., a distributed denial of service or DDOS attack) .Training models on benign and malicious samples to help them predict whether new samples are malicious Reinforcement machine learning is a model used for machine learning in cyber security that is similar to supervised machine learning. However, reinforcement machine learning trains the algorithm by trial and error rather than using sample data. Positive or negative cues are given and registered along the way, with the algorithm programmed to seek affirmation and avoid penalties.

Unsupervised machine learning in cyber security is used to analyze and cluster unlabeled datasets (e.g., photo images, audio and video recordings, articles, or social media posts). It can identify hidden patterns or data groupings without human intervention.  The algorithm scans through data sets, looking for patterns that are used to group information into subsets. Unsupervised machine learning is most commonly used for deep learning. Benefits of machine learning in cyber security. Enables BYOD (bring your own device) and CYOD (choose your own device) to be securely implemented

Detects threats in the early stages. Enables adaptable and proactive defense systems Identifies hard-to-find network vulnerabilities. Internalizes learning from previous attacks to prevent future attacks based on similar profiles. Makes it easier for security analysts to quickly identify, prioritize, and remediate attacks. Minimizes human errors and Powers sophisticated authentication mechanisms, such as facial recognition, fingerprint recognition, motion tracking, retinal scanners, and voice recognition. Helps prevent security threats against endpoints provides insights into advanced threats.

*Future Work*

The development of machine learning-based cyber security techniques for detecting upcoming cyber attacks is a dynamic and evolving field. Future work can focus on several key areas to enhance the effectiveness, scalability, and adaptability of these systems: Explore and implement more sophisticated machine learning models, including deep learning architectures such as convolution neural networks (CNNs) and recurrent neural networks (RNNs), to improve the accuracy and robustness of threat detection. Investigate the use of ensemble learning techniques to combine multiple models and improve overall detection performance.

Optimize algorithms for real-time processing to ensure timely detection and response to cyber threats. This includes reducing latency in data collection, feature extraction, and model inference. Develop lightweight models that can be deployed on edge devices to enable real-time threat detection closer to the source. Implement mechanisms for continuous learning and adaptation to new threats. This includes online learning techniques that allow the system to update its knowledge base in real-time as new data becomes available. Explore reinforcement learning approaches to enable the system to adapt its defense strategies based on evolving threat landscapes. Integrate the machine learning-based detection system with existing threat intelligence platforms to leverage shared threat information and enhance the system's ability to detect and respond to emerging threats. Develop APIs and interoperability standards to facilitate seamless integration with other cyber security tools and platforms. Focus on scalability to handle large-scale data environments and high-velocity data streams. This includes leveraging distributed computing frameworks such as Apache Kafka and Apache Spark. Explore cloud-based solutions to provide scalable and flexible deployment options that can accommodate varying workloads. Enhance feature engineering techniques to extract more relevant and discriminative features from raw network data. This includes incorporating domain knowledge and leveraging advanced data enrichment methods. Utilize external data sources, such as dark web monitoring and social media feeds, to enrich the data and improve the context for threat detection. Develop methods to increase the explain ability and transparency of machine learning models. This is crucial for building trust with security analysts and stakeholders, and for meeting regulatory requirements. Implement explainable AI (XAI) techniques to provide clear and interpretable insights into the model's decision-making process.

Investigate techniques to enhance the robustness of machine learning models against adversarial attacks. This includes developing defenses against evasion and poisoning attacks that attempt to deceive or corrupt the model. Conduct adversarial testing and implement countermeasures to ensure the resilience of the system. Integrate user and behavioral analytics to detect insider threats and account compromise activities. This involves analyzing user behavior patterns and detecting deviations from normal behavior. Develop models that can correlate user activities with network events to provide a comprehensive view of potential threats. Ensure that the system complies with relevant regulations and standards for data privacy and cyber security. This includes adhering to guidelines from regulatory bodies such as GDPR, HIPAA, and NIST. Address ethical considerations related to the use of machine learning in cyber security, including data privacy, algorithmic fairness, and the potential impact on employment. By focusing on these areas, future developments can significantly enhance the capability, reliability, and user acceptance of machine learning-based cyber security techniques, ultimately leading to more effective and proactive defense against cyber attacks.

## Conclusion

This study meticulously evaluates the Random Forest algorithm for intrusion detection within autonomous vehicle systems, focusing on its efficacy in accurately distinguishing between attack and non-attack network traffic. The algorithm's high precision in identifying legitimate traffic is commendable, showcasing 93.6 % accuracy for non-attack packets. However, the detection of attack packets reveals room for improvement, with an 87.6 % accuracy indicating a propensity for false negatives. Machine learning (ML) has revolutionized the field of cyber security, providing powerful tools for detecting and mitigating upcoming cyber attacks. By leveraging vast amounts of data and advanced algorithms, Machine learning -based cyber security techniques offer a proactive approach to threat detection and defense. The key advantages of Machine learning in cyber security include its ability to Detect Anomalies Machine learning algorithms can identify unusual patterns and behaviors that may indicate a cyber attack. Techniques such as anomaly detection are crucial for recognizing deviations from normal network activity, which are often early signs of malicious activities. Predict Threats Predictive analytics, powered by Machine learning, can forecast potential security threats by analyzing historical data and identifying trends. This capability allows organizations to anticipate and prepare for possible attacks before they occur. Machine learning enables automated threat detection and response, reducing the time and effort required for manual analysis. Automated systems can quickly respond to threats, minimizing potential damage and ensuring a faster resolution.

## Reference

[1]. Dipankar Dasgupta. Immunity-based intrusion detection system: A general framework. In Proceedings of the 22nd National.
[2]. Systems Security Confer-ence (NISSC). Arlington, Virginia, USA, 1999.
[3]. https://www.sciencedirect.com/science/article/abs/ pii/S03608352292#:~ :text=I n% 20t he %C%20different%20algorithms,for%20threat%20identification%20and%20defense.
[4]. Jonatan Gomez and DipankarDasgupta. Evolving fuzzy classi_ers for intrusion detection.
[5]. In Proceedings of the 2002 IEEE Workshop on Information Assurance,West Point, NY, USA, 2002.
[6]. Steven A. Hofmeyr, Stephanie Forrest, and Anil Somayaji. Intrusion detection using sequences of system calls.
[7]. A. Shabtai, E. Menahem and Y. Elovici. FSign: automatic, function-based signature generation for malware.
[8]. https://ieeexplore.ieee.org/document/101927.
[9]. K. Graves, Ceh: Official certified ethical hacker review guide.
[10]. R. Christopher, "Port scanning techniques and the defense against them," SANS Institute, 2001.
[11[.S. Robertson, E. V. Siegel, M. Miller, and S. J. Stolfo, "Surveillance detection in high bandwidth environments,"

[12]. M. Thangamani, and Jafar Ali Ibrahim. S, "Knowledge Exploration in Image Text Data using Data Hiding Scheme.