Noah Philippe 14398743

Aleksyah Vien 24488013

Jemii Virakorn 24506875

Megan Siow 14049427

Machine Learning Attacks: Securing Wireless Systems

**November 1, 2024**

# 

# 

# Research Question

What are the hidden vulnerabilities in machine learning (ML) integrated wireless systems, and how can these weaknesses be addressed to develop more robust and secure systems?

# Research Project Goals

Our project aims to raise awareness about the vulnerabilities and weaknesses of ML-based wireless systems by examining adversarial attacks and data poisoning while providing practical solutions to mitigate these risks. We will compare relevant research papers on machine learning and wireless system topics to understand how these threats affect network security and performance. The research will also contain a practical demonstration of these vulnerabilities impacting ML-integrated wireless systems in a MATLAB example to show the real-world effects of these vulnerabilities and provide an in-depth explanation of the dangers involved.

[**Research Question 1**](#_b89khlia5ymn)

[**Research Project Goals 1**](#_inkfeeqt261)

[Abstract 4](#_65tpvkxeyx83)

[**SECTION 1: Literature Review 5**](#_idjuzupo18qq)

[Introduction 5](#_7c9zii8mua4u)

[Topic 1: Adversarial attacks 6](#_myh1pbjqge20)

[Topic 2: Data Poisoning and Data Pollution on Machine Learning 8](#_ebe0r5l57oj0)

[Topic 3: Machine Learning Vulnerabilities in Signal Classification 10](#_j80eyqsqfinl)

[Gaps in Research 12](#_omuwelw6lxab)

[Conclusion 13](#_bu1j4dbfxa32)

[**SECTION 2: Research Report 14**](#_z20hzyikl8cc)

[Executive Summary 14](#_781gjpmqoqbr)

[Introduction 14](#_4bdtdwkyfrwt)

[Methodology 15](#_dvsbtql9bw8t)

[MATLAB Simulation 15](#_ne2anfdfpxb6)

[Findings 18](#_6b0jrslbsrjt)

[Adversarial noise performance 18](#_w62rm6bu0vyg)

[Data Poisoning 19](#_wvy4imjy12sc)

[Anomaly Detection 20](#_8rdmkm6xa5bk)

[Discussion 21](#_ep5wtai57c74)

[Recommendations 23](#_3y00gpdw4ei8)

[Conclusion 26](#_vpwr0jp9k7ly)

[**Appendix 28**](#_8rxh7i8w8qu2)

[A1. Contributions table 28](#_dohjjqsf22ko)

[**References 29**](#_m63inyaoj7gf)

# 

## 

# Abstract

Integrating Machine Learning (ML) into wireless communication systems has significantly enhanced signal classification, anomaly detection, and overall efficiency. However, technological advances also bring new vulnerabilities; adversarial attacks present a growing security threat. This paper investigates these attacks, focusing on the impact of data poisoning and noise interference. The research completed through the literature review and a practical simulation in MATLAB explores the accuracy and robustness of ML models in these systems. The simulation focuses on generating a wireless signal and introducing adversarial noise. It will apply data poisoning to test the performance of traditional and deep learning models like Support Vector Machines (SVM) and a basic neural network. Using performance metrics like F1-score, recall, and checksum revealed a decline in classification accuracy, with data poisoning being a significant threat. The results suggest that further development of security mechanisms is necessary to mitigate adversarial influences more effectively, underscoring the need for more resilient models and robust systems.

## 

# 

# 

# SECTION 1: Literature Review

## **Introduction**

After reaching a growth target of $1.58 billion in 2017, machine learning (ML) is expected to reach a

growth target of $20.83 billion in 2024 (Chen et al., 2021). The MLs' incorporation improves the wireless systems' performance, efficiency, and adaptability, creating a greater reliance on ML-based models (Soori et al., 2023). It quickly processes large amounts of data, enhances signal modulation and classification, and uses patterns found in the data to make predictions (Sun et al., 2019).

However, the dependability of ML in these wireless systems makes them more susceptible to adversarial

attacks where attackers can manipulate data to create confusion and exploit ML's flaws. These attacks take advantage of hidden vulnerabilities in ML models which can result in incorrect classifications and crippled system performanc which can affect ML decision-making and patterns. Poor network performance or incorrect signal classification may result from inadequate defences against these attacks (Apnic Foundation, 2022). Therefore, it is vital to ensure wireless system security.

Existing research provide insight into ML impact on wireless system and their vulnerabilities to

adversarial attacks and this literature review examines critical studies on adversarial attacks, data poisoning and vulnerabilities in signal classification. It will highlight gaps in the existing research, identify areas where future work is needed, and emphasise the need for more robust defence mechanisms for the system to withstand adversarial attacks in unpredictable environments. Future research is necessary to create a more resilient wireless system capable of withstanding attacks better.

This raises the question, ‘What are the hidden vulnerabilities in ML-integrated wireless systems, and how

can these weaknesses be addressed to develop more robust and secure systems?’ This question allows the review to comprehensively analyse current challenges, showing the importance of future research on developing solutions for a more robust ML system. This is to ensure the long-term reliability and security of ML-powered wireless systems.

## **Topic 1: Adversarial attacks**

Within recent years, several studies have highlighted the growing concern about the hidden

vulnerabilities of ML wireless systems, specifically concerning adversarial attacks. Adversarial attacks are techniques specifically designed to deceive, exploit and interrupt the performance of ML wireless systems, causing them to generate incorrect predictions (Boesch, 2023). Researchers such as Goodfellow et al (2014) were among the first to introduce the concept of adversarial attacks. Their research discovered the existence of adversarial attacks and that even the most minor amounts of noise could perturb the most optimised ML models. Proceeding with this foundational work, research has further shown that adversarial attacks have evolved where small amounts of perturbations can manipulate images to a degree where the system misclassifies them as a completely different object (Hebbar, 2023). These adversarial attacks highlight the fragility of ML wireless systems, providing incorrect data from manipulated input noise.

Due to the nature of modern wireless systems, specifically those that utilise 5G technologies or

Internet-of-Things (IoT) ecosystems, the threat of adversarial attacks increases. Studies completed by researchers such as Adesina et al (2022) detail how adversarial attacks continue to evolve as wireless communication systems continue to develop. The study highlights how these attacks have evolved from sending jamming signals to overwhelm the communication system to becoming adversaries that send out small perturbation signals to disrupt wireless communication. The difference between the two signals is identified through how they affect the system, with jamming signals completely disabling communication. In contrast, permutation, or interference signals, manipulate the data to provide incorrect outcomes. According to researchers such as Usama et al (2020), some adversarial attacks can successfully target vulnerable ML wireless systems without prior knowledge of the system, known as a black-box attack. These black-box attacks demonstrate the delicacy of ML wireless systems, spotlighting that all systems are highly capable of being attacked, calling for the desperate need for further research to prevent such attacks.

Through the advancements made in understanding the impact of adversarial attacks, researchers

have been able to place their focus on developing strategies to mitigate attacks and preserve the integrity of ML wireless systems. Studies indicate how Adversarial Machine Learning (AML) has been designed to mitigate attacks and highlight the need to enhance ML wireless systems' robustness. Researchers such as Alhajjar et al. (2021) present the study of AML as an effective solution to identify and understand different adversarial attack strategies and propose defensive solutions to improve the robustness and adaptability of ML wireless systems against adversarial attacks. Attacks explored within the paper include exploratory, evasion, causative and trojan adversarial attacks. The proposed defence solutions include adversarial perturbations based training of the ML system, autoencoder based defence against the attacks at component level, and obfuscation of the transmission signal and distract the intruder. Despite their ability to increase the resiliency of the ML wireless system, these types of solutions however would necessitate an ever-evolving adaptation to safeguard the system from the constantly evolving adversarial attacks (Macas et al, 2024). It indicates that more robust solutions must be developed that can accommodate the competitive demand to enhance the protection profile of an ML wireless system.

While ML integrated wireless systems may provide numerous benefits, the hidden vulnerabilities must

be addressed. Further research should be conducted to fully comprehend the complexity of adversarial attacks and allow for the development of a more reliable and secure ML wireless system.

## 

## **Topic 2: Data Poisoning and Data Pollution on Machine Learning**

Machine learning models are vulnerable to exploitation via data poisoning attacks. After the training

phase, adversaries introduce faulty or biassed data into the training dataset, leading the model to learn erroneous patterns later exploited after deployment; such assaults might occur. Data poisoning compromises the model inherently, lowering its capacity to provide accurate classifications, unlike adversarial assaults that change inputs during the prediction phase. Misclassifications can cause significant security flaws in wireless systems, making this problem more dangerous (Jagielski et al., 2018). Data poisoning threats have been thoroughly examined in broad machine learning applications; nevertheless, research focused on wireless systems remains somewhat limited, with most research focusing on natural language processing or image classification, ignoring the specific problems of signal classification in wireless situations (Steinhardt et al., 2017). This limitation emphasises the importance of domain-specific research suited to wireless communication networks' unique characteristics. Although data poisoning vulnerabilities have been seen in several domains, wireless applications provide specific issues due to their dynamic nature, including the possibility of service disruptions (Bhagoji et al., 2019). Moreover, the role of data poisoning attacks in modulating communication signals still needs to be explored, as suggested by Liu et al. (2023), who argue for more focused research in wireless communication.

Data poisoning attacks may be extremely harmful in distributed systems such as federated learning,

which uses data from several devices to build a global model. Because Federated Learning is distributed, it is easy for attackers to import compromised data from a compromised device, affecting the integrity of the global model. According to Bagdasaryan et al. (2020), this decentralised structure makes federated systems especially susceptible since poisoned changes to local models can compound to have a significant effect on the global model. Wireless networks' dynamic and real-time nature makes it much more difficult to build data poisoning remedies. Most current research assumes offline training settings where models are developed and deployed display a lack of understanding concerning the effects of poisoning attacks on real-time systems and the implementation of protections without causing performance-degrading delays. In environments such as wireless communication systems, this lack of adaptation to real-time training poses critical risks (Wu et al., 2020). Furthermore, it is getting increasingly difficult to distinguish between intrinsic noise and data corruption in wireless systems susceptible to much interference and noise (Ma et al., 2022). How effective mitigation strategies like data validation and differential privacy are in high-noise contexts is uncertain. The following research should aim to develop noise- and environment-resistant defences that can tolerate substantial degradation of environmental signals. This concern is reinforced by research by Goldblum et al. (2020), which shows that wireless systems are especially vulnerable to attacks since current countermeasures frequently fail to generalise in noisy environments.

Another critical area that needs to be examined is the lasting influence of data pollution on machine

learning models. While a large body of research focuses on the immediate performance decline, more is needed about how long it takes for damaged models to recover and if retraining without extra safeguards is effective (Xiao et al., 2015). Long-term machine learning (ML) systems deployments are essential in wireless networks because prolonged periods of reduced performance can lead to catastrophic failures, particularly in time-sensitive systems like emergency response or defence networks (Wu et al., 2020).

## 

## **Topic 3: Machine learning vulnerabilities in signal classification**

The ability to classify signals has been dramatically improved by integrating ML into wireless systems,

providing flexibility and scalability. ML uses a signal database with signals collected under different conditions, which it then uses as a base to analyse the unique characteristics of a signal to distinguish between signal types. This database is divided into two parts; a test set used to assess the model performance and a training set used to teach it to aid the model's ability to develop precise predictions (Nova Systems, 2023). However, the investigation revealed severe flaws in these systems, raising questions about their security and dependability.

Signal misclassifications resulting from data bias can appear during the training phase of an ML model.

"Data bias" occurs if there are incomplete or incorrect training sets. This could lead to skewed predictions, negatively impacting decision-making and system performance (Pragmatic Institute, 2024). Mehrabi et al. (2021) clarify that ML systems may reflect bias or discriminatory behaviour in their training data and may be introduced during data collection. As these systems are frequently used to make crucial decisions, there must be a strong awareness of biases; otherwise, it can result in skewed predictions or compromised system performance. This is similar to Zhang et al. (2019), who discuss how bias can be introduced during training phases through data injection, modification, or logic corruption. Incorrect data that is sent into the system can affect its performance. For example, when an ML system is trained with altered data, it could mistakenly classify unaltered data as an attack because it sees it as an “outlier” from what it was taught. Both papers state how data collection and preparation are critical steps in ML models. This is because errors in this stage can further lead to significant mistakes.

Attack on the physical layer can result in misidentifying modulation schemes. Modulation schemes are

methods used to encode data into wireless signal for transmission critical for accurate communication. Therefore misidentifying this could mean the system misinterprets signals or fails recognise the correct pattern of signals (Reis, 2024). Khazane et al. (2024) say the physical layer is vulnerable to security threats such as inserting fake data, malicious code, and sniffing or snooping. There was an experiment on graph neural network (GNN) models, which effectively classify signals and identify modulation schemes based on the relationships and patterns within data. The classification precision decreased by over 30% when exposed to adversarial attacks making the system unreliable because it could lead to incorrect predictions and classifications. They state that improving the attack resistance of modulation classification is crucial for enhancing the security of the physical layer. This is similar to Freitas (2023), who discusses VT-CNN2 modulation classifiers, a popular ML technique that identifies the method to encode a signal for transmission. The classifier identifies the modular scheme, and depending on the signal-to-noise ratio (SNR), the model's classification performance varies. Significant noise in a lower SNR causes data perturbations making classification more challenging for the model. Due to these attacks, the model may incorrectly identify the modulation scheme, resulting in improper signal processing. Communication may become compromised as a consequence. ML in these systems has proven to provide good classification results; however, it raises questions about security and reliability.

Wireless communication is transmitted over the air and shared publicly so attackers can tamper with

signals more easily. Although ML techniques present several benefits for wireless systems signal classification, the articles also highlighted potential vulnerabilities. Future studies must, therefore, concentrate on ways to lower signal misclassification, suggesting more robust ML models and strong defences against security threats.

## 

## **Gaps in Research**

ML continuously enhances signal classification and modulation in wireless communication systems;

However, the growing threat of adversarial attacks highlights a gap in current defence systems. Most research is conducted in a controlled environment with minimal interference, which is different from reality, where the environment is more unpredictable and noisy, particularly in 5G or IoT systems. Present adversarial defence strategies, like anomaly detection and adversarial training, find it difficult to remain effective in real-world applications. This raises the question: What are the hidden vulnerabilities in ML-integrated wireless systems, and how can these weaknesses be addressed to develop more robust and secure systems?

Those factors reduce the efficiency of current defences, so there is a need to research noise and

environment-resistance defences. Investigating and enhancing strategies to ensure system resilience will fortify these systems against adversarial attacks so there are no delays, performance degradation, signal misclassification, and reduced chance of data poisoning. By addressing these gaps in the review, hidden vulnerabilities can be identified. To mitigate these weaknesses, ongoing research must concentrate on how a more robust and improved ML model could be created to withstand adversarial attacks even in a degraded environment.

## 

## **Conclusion**

ML continues to integrate into wireless systems, significantly improving the technology. However, it also

creates more vulnerabilities.. The review of current research raises important questions about data poisoning, adversarial attacks, and how attacks may impact signal classification. Although ML has improved wireless systems' signal classification and modulation, there is still a weakness in the current defence mechanisms because of the unpredictability of the environment. Defence systems susceptible to high levels of interference, like adversarial training and anomaly detection, are not robust enough.

As a result, more robust defences that can adjust to the complexity of today’s wireless systems must be

developed. Developing noise-resistant defences that preserve system performance must be the primary goal of future research. The security and dependability of wireless systems can be improved by creating more robust ML models and improving defence techniques to guarantee their durability in the most demanding settings.

# SECTION 2: Research Report

## **Executive Summary**

This report aims to evaluate and address the vulnerabilities that affect Machine Learning (ML) integrated

wireless systems and provide actionable defensive recommendations to enhance its efficiency and reliability. Studies have shown that while the reliability of these wireless communication systems have improved significantly, the vulnerabilities that affect these systems have also advanced extensively.

In the practical demonstration of this report, Matlab will be used to simulate a wireless communication

system, and the attacks will affect the wireless signals. The models used to simulate the ML system were Neural Networks and Support Vector Machines, and the attacks will simulate adversarial attacks, such as noise interference and data poisoning attacks. From this demonstration, several conclusions have been drawn.

1. When adversarial noise is added to a clean wireless signal, the model can detect these anomalies with a precision score of 0.4-0.6.
2. When data poisoning attacks were added to a clean signal, the model could detect the anomalies with a 50%-80% accuracy rating.

These findings indicate an ML wireless system’s inability to detect and prevent different types of attacks

effectively. The simulation further indicated that the system would generate false reports of an adversarial attack.

To address these vulnerabilities, this report recommends several mitigation strategies. These defensive

strategies include:

* Autoencoders which are used to detect and filter unrecognisable signals
* Noise-resistant algorithms which prevent the influence of noise on a wireless model
* Continual learning frameworks that allow the model to constantly evolve based on any changes in its environment.

Developing a secure and attack-resilient ML integrated system can be undertaken by addressing the

vulnerabilities within an ML wireless communication system and the attacks it is susceptible to.

## **Introduction**

**Research Question**

The rapid rollout of ML into wireless systems has introduced many benefits to wireless transmissions such as advanced adaptability and greater operational efficiency. ML is responsible for signal classification and anomaly detection within wireless systems, helping optimise wireless transmissions and communication systems. The integration of ML algorithms into wireless systems also brings out several challenges in the operations as learning methods can easily be exploited by bad actors. These adversarial entities may then conduct malicious actions including signal jamming, data poisoning and adversarial noise to influence data (or model predictions) to threaten the security and stability of the network.

**Introduction to the Project Development Process**

This study aims to address the vulnerabilities affecting ML-integrated wireless systems through

both an analytical literature review and a practical demonstration. The practical demonstration will utilise MATLAB to simulate a wireless signal transmission influenced by adversarial attacks, including noise interference and data poisoning. The simulated wireless network will assess the impacts of these attacks on machine learning models. The research aims to identify gaps in existing research studies and ML mechanisms, provide possible strategies to mitigate risks and heighten the security and robustness of machine learning integrated wireless systems against advancing threats.

## 

## **Methodology**

This study uses a dual approach, including a literature review and practical simulation in MATLAB. The

goal is to produce insight and detailed analysis of vulnerabilities in ML wireless systems and demonstrate the practical impact of adversarial attacks.

### MATLAB Simulation

The primary purpose of this MATLAB project is to simulate the transmission of wireless signals in a

communication system and highlight the vulnerabilities associated with machine learning and its effect on the integrated system. Our MATLAB project focuses on machine learning performance benchmarking using f1 score, recall and checksum value to evaluate the model's accuracy.

These performance matrix values allow us to draw the impact of malicious attacks against the model,

such as data poisoning and adversarial noise. This project aims to analyse the performance of machine learning algorithms in the event of these attacks occurring in a wireless communication environment. Traditional and deep learning methods will be used to classify wireless signals using a Neural Network with three hidden layers and a Supervised Vector Machine to evaluate the impact of ML techniques on wireless communication.

The primary machine learning model used is Support Vector Machine Classification for assessing

machine learning weaknesses. Our group chose this method due to SVM’s robustness and accuracy in handling small datasets and its functionality of defining clear boundaries. Wireless communication systems often involve classification where clean signals have to be separated from noisy environments, and SVM’s effective classification makes it ideal for making judgements.

Furthermore, deep learning will be examined with a simple neural network featuring a three-layer

architecture. The neural network will classify clean and adversarial signals while assessing accuracy against data poisoning, similar to the SVM model. While various research articles found Recurrent Neural Networks (RNN) to be robust in handling time-series data similar to wireless signals, this deep learning technique has certain restrictions. RNNs are prone to overfitting when trained on small datasets and require fine-tuning of hyperparameters. The project assesses the vulnerability of machine learning models in wireless systems, which involves classification tasks with efficient basic anomaly detection techniques, and simple neural networks are used to train both the anomaly detection system and classification for clean and adversarial signals.

1. **Wireless Signal Generation**

### 

The first section generates a simple wireless signal to simulate a real-world wireless situation. The signal created is a 5 Hz frequency sine wave representing the communication signal. After this process, random noise is introduced and alters the signal, creating an adversarial signal that simulates corruption and adversarial attacks that threaten machine learning algorithms.

1. **Adversarial Attack Simulation**

### 

This section trains a Support Vector Machine model to distinguish between clean and adversarial signals, being trained on the first modules’ signal datasets and evaluating the models’ ability to detect adversarial noise. The dataset consists of the combined clean and adversarial signals, labelled 1 and 0, respectively, allowing the SVM and NN models to define clear decision boundaries and classify the signal. Finally, the model evaluates the classification results with precision, recall and F1-score with checksum.

1. **Data Poisoning**

### 

In this module, we assess the models’ resistance to data poisoning attacks; this is accomplished by changing the value of clean signal labels to emulate the effects of data poisoning. The code intentionally poisons 20% of the dataset, flipping bits from 1 to 0 and retraining the support vector machine model based on the poisoned dataset. The model's predictions are then calculated, producing performance matrix scores such as F1- score, accuracy and checksum, which evaluates how corrupted data impacts this model’s accuracy.

1. **Anomaly Detection**

Anomaly detection is performed with the Support Vector Machine model, using a noisy signal, in order to detect and classify unexpected modifications and distortions from the normal signal transmission data (wireless). Demonstrating how machine learning can be employed as a strength in wireless communication networks to detect likely anomalous signal behaviour

### 

## 

## **Findings**

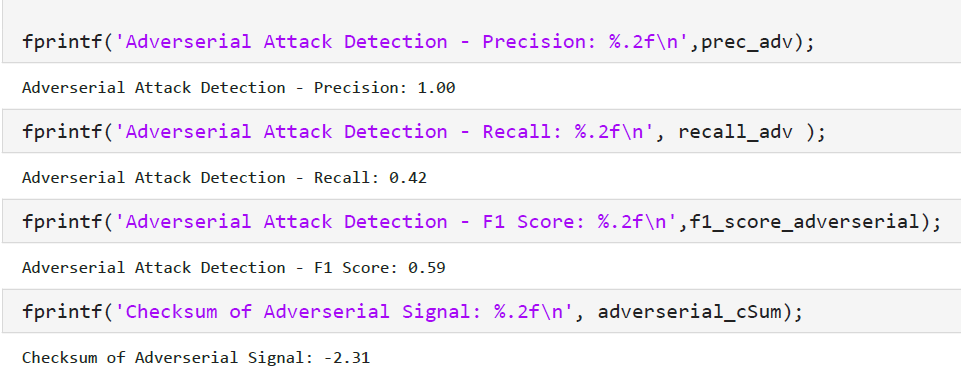
In the MATLAB demonstration, we closely evaluated the impacts of data poisoning and adversarial noise

attacks on machine learning algorithms assessing Neural Network (NN) and Support Vector Machines (SVM) model accuracy and resilience to changed data inputs.

### 

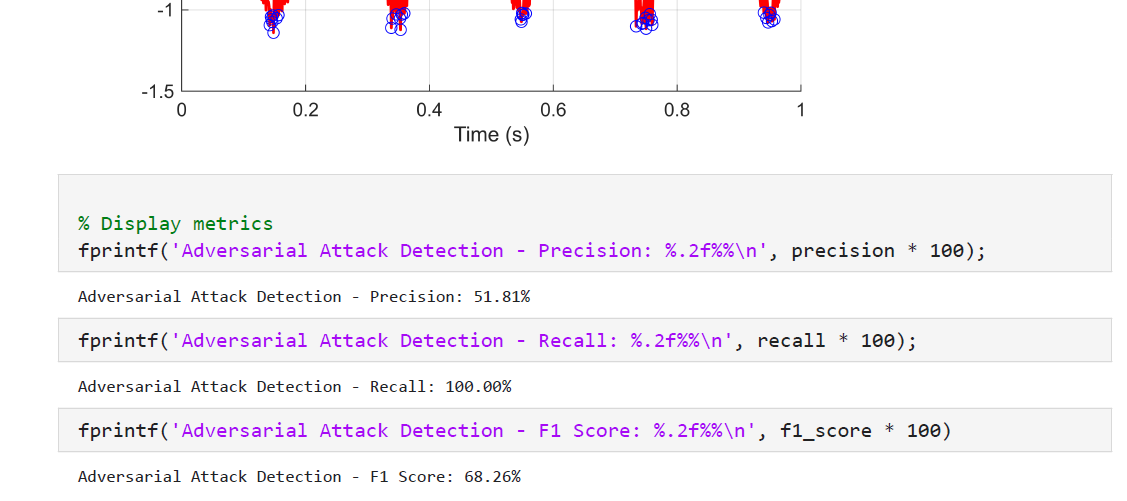
### Adversarial noise performance

Adversarial noise was added to clean signals in SVM and NN models; the SVM Model achieved a higher precision score of 1.00; however, it had trouble detecting adversarial inputs with a recall score of 0.42. This highlights how the model has difficulties detecting all adversarial cases but can accurately identify most adversarial samples.



*Figure 1: Adversarial Attack Detection of SVM model*

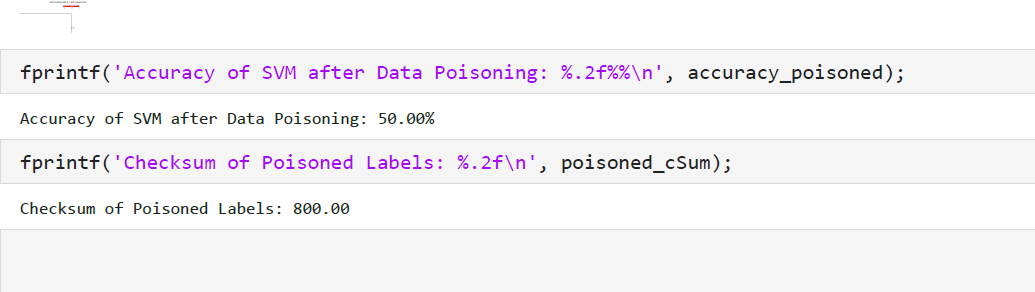
The Neural Network exhibits a high recall score of 1.00 with a lower precision of around 0.52; the model can identify more anomalies; however, this results in some of these classifications being false positives. Deep learning allows the neural network model to achieve a higher f1 score overall, balancing precision and recall better than the SVM model. The NN model is more suitable for detecting adversarial cases than support vector machines.



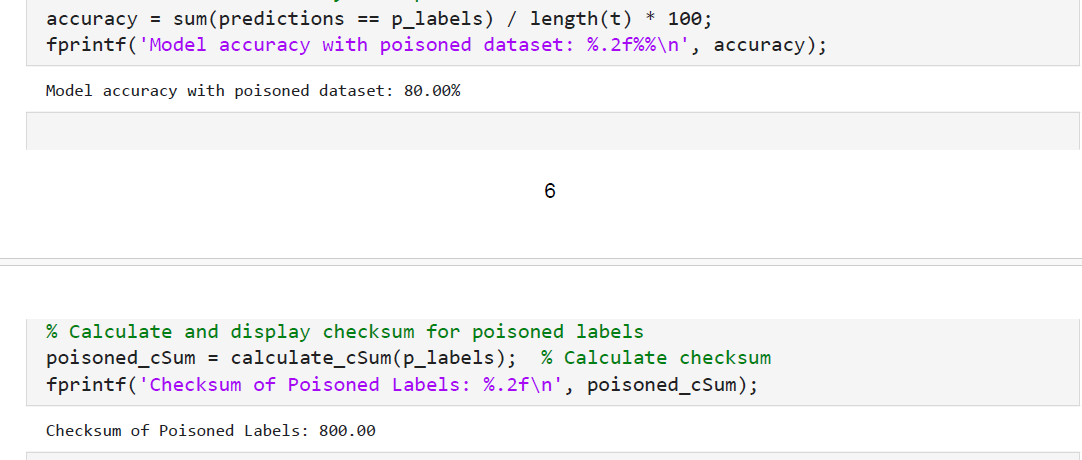
*Figure 2: Adversarial Attack detection*

### Data Poisoning

A data poisoning attack takes place, when we enter corrupt or faulty data into the training dataset. They were then interrupting the process of machine learning. As shown in Figure 3, it created a 50.00% decrease in accuracy on SVM model after data poisoning assault in our study, which indicated that the model capability could be compromised with flipped signal bits. The checksum value has proliferated, confirming the departures from the model's expected results.



*Figure 3: Data Poisoning attack in SVM model*



*Figure 4: Data Poisoning attack results with Neural Network model*

In opposition to this, the SVM outperformed the Neural network on the data poisoning simulation test. As we can see in Figure 4, the model presents an 80% accuracy score, indicating that the NN model is more resistant to label changes and poison dataset. The designs can search complex relationships and patterns in the data, as this is a layered design which has a nonlinear boundary due to neural networks. Hence, compared to SVM, neural networks are able to withstand an input that has a flipped label since this model is focused on recognizing linear interplays.

### Anomaly Detection

The SVM model was deployed to find anomalies in wireless signals where the model matches the signal of a “normal” archetype during the training phase (Figure 5). The model fails to classify all the anomalies correctly, resulting in missing anomalies or classifying clean data as anomalies. The inconsistency of SVM in anomaly detection is mainly due to static thresholding, since the real-world noise changes from time to time and the SVM model does not change accordingly to that change. Dynamic thresholding can eventually be used to increase accuracy in distinguishing between anomalies and high variations in wireless signals.

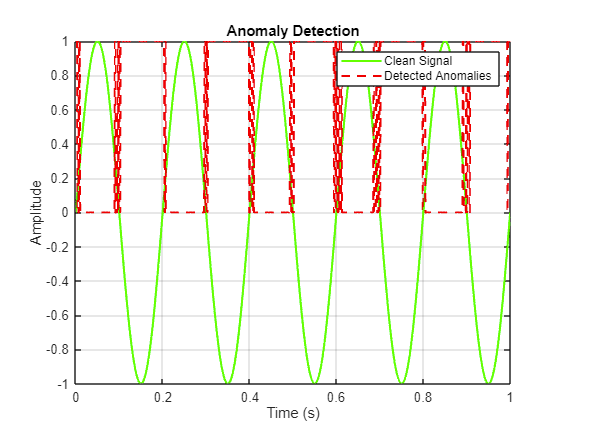


Figure 5 Anomaly detection with SVM

This practical description discusses how; adversary noise and data manipulation contributes to the overall integrity and reliability of machine learning based models. Adversarial noise during inference severely affects the decisions of the model however data poisoning has an apparent influence on the accuracy of the training data which leads to incorrect classifications and thus to a loss of learning in the model.While both attacks negatively affect model performance, data poisoning ultimately is the greater risk because data is changed during the training phase, which makes it more difficult to classify even clean signals.

## 

## **Discussion**

This project aims to assess how challenges and vulnerabilities affect the security and performance of ML-integrated wireless systems. The MATLAB simulation highlights how adversarial attacks compromise these systems reducing the overall validity of the network. The findings will contribute to developing more resilient defences against attacks and threats. Research on creating more secure ML-integrated wireless systems could be discovered by identifying key issues and challenges.

ML algorithms can be exploited, impacting network stability and reliability. ML models are responsible for classifying wireless signals and are susceptible to adversarial attacks like signal jamming. These attacks manipulate the signal’s data, distorting the model’s ability to classify signals correctly. They distort it by introducing malicious data in the model’s training set, which will corrupt its learning proces (Jagielski et al., 2018). This can lead to poor performance, miscommunication, and false negatives and positives. A false positive occurs when a model detects an attack when nothing is wrong, resulting in incorrect decision-making and wasted resources, while a false negative is when it fails to detect a real threat. Both can significantly weaken the security and stability of the wireless system. Attacks like noise interference potentially destabilise the system lowering its operational efficiency and jeopardising the integrity of the network (Rahul, 2023). A less robust ML model reduce the security and performance of wireless transmissions.

There is a weakness in the defences currently in place against adversarial attacks. The evolving nature of these threats makes it challenging to develop strategies to cover all aspects adequately. However, it could be adapted to reduce the amount of current vulnerabilities. As discussed in the literature review and the MATLAB simulation, anomaly detection and adversarial training strategies struggle in more noisy and unpredictable environments. The interference during transmission directly influence accuracy of the ML model’s classification ability. These systems have hidden vulnerabilities identified in the literature review, which were exposed by the MATLAB simulation testing the system's strength when under attack. This simulation provided a practical demonstration of how noise can affect model performance.

This study explored both traditional and deep learning ML models to evaluate their effectiveness against adversarial attacks in wireless communication systems. The Support Vector Machine (SVM) was selected because it can model complex nonlinear decision boundaries suited for classifying clean and adversarial signals in a controlled environment (Cosma et al., 2017). However it struggles with handling complex adversarial noise and large datasets, underscoring the need for adaptable models capable of handling a more unpredictable environment (Dhiraj, 2019). This contrast with a simple neural network with a three-layered architecture representing a deep learning approach. It is more likely to learn from harmful pattern in poisoned data, leading to overfitting. The model could adapt too closely to the manipulated training data (Vignesh, 2020). This study's SVM and neural networks illustrate the strengths and weaknesses of traditional and deep learning models. Therefore, advanced architectures or hybrid models are needed to achieve classification accuracy and resilience in dynamic wireless networks.

One of the most pressing challenges is the development of ML models that can remain robust in noisy environments. The MATLAB simulation, noise interference was introduced to show how the current models degrade in performance, especially regarding signal classification and anomaly detection. Therefore, there is a need for more noise-resistant strategies, which is a significant gap in existing research. The literature review also highlighted the impact of adversarial attacks, such as signal misclassification and performance degradation, further confirmed in the simulation; adversarial noise introduced significant classification errors (Freitas, 2023). Therefore, there is an issue of how to balance security and performance. Also, with the variety of wireless systems, it is challenging to design a solution that fits them all, such as 5G and IoT.

## 

## **Recommendations**

With increased dependency on machine learning, especially in wireless communication systems,

performance, efficiency, and adaptability are expected to see unprecedented application development in 5G technology and IoT (Chen et al., 2021; Soori et al., 2023). Integration of machine learning allows for fast data processing, enhancements in signal modulation, and exploitation of patterns in data to get better prediction and classification (Sun et al., 2019). The extended role of ML in wireless networks further opens new doors for vulnerabilities, hence exposing these systems to a wide range of security threats such as adversarial attacks, data poisonings, and misclassifications of signals (Jagielski et al., 2018; Mehrabi et al., 2021). It is important to note that the latent weaknesses of the ML models are leveraged in such forms of attacks, leading to potential disruptions in network performance, information integrity, and general system security.

Given these risks, solid defences and adaptive security mechanisms must form an intrinsic part of

protection for ML-integrated wireless systems against adversarial influences. While existing literature has pointed to the threats, studies show that there is still a lacuna in research that presents real-world solutions, particularly in dynamic and noisy environments (Wu et al., 2020; Macas et al., 2024). Addressing such vulnerabilities is of prime importance to retain wireless network reliability while continuing to evolve. The paper discusses the current challenges, identifies the critical research gaps, and suggests recommendations that would help reinforce security and resilience in ML-powered wireless systems to continuously make them effective and trustworthy in the face of upcoming threats.

The adversarial attack is a massive threat in ML-based wireless systems, where active attackers can

modify the signal to mislead models and degrade system performance (Boesch, 2023). Advanced adversarial training may be enabled by training the model with various types of disruptions and various levels of adversarial attacks. Therefore, the models become resilient against these manipulations since they "learn" from the various ranges of adversarial input. Advanced techniques employ dynamic training methods that help the models adapt to emerging and various attack strategies (Alhajjar et al., 2021). This is enabled by integrating real-time feedback, which helps cope with emerging threats inside a noisy and unpredictable wireless environment, such as those found in 5G and IoT systems (Adesina et al., 2022).

The data poisoning attacks can scale to the federated learning systems where data is distributed across

devices; an attacker may introduce poisoned data into one or more local devices. Federated learning architectures have enhanced data validation mechanisms to protect against such data poisoning (Bagdasaryan et al., 2020). One of the efficient approaches is differential privacy, a technique that adds noise in a controlled manner to guarantee privacy protection. This reduces the impact of possible poisoned data, resulting in challenging conditions for poisoned inputs to alter the global model (Liu et al., 2023). This will provide resistance against malicious alteration of the federated learning model and enhance the security of a wireless communication network. Further research may also focus on hybrid methods of validation – a combination of traditional data validation and ML-based anomaly detection models to offer real-time poisoning detection.

ML wireless systems face severe problems with noisy environments, leading to misclassifications in

signal modulation schemes and weakened system reliability. Noise-resistant algorithms can then be developed based on these approaches using signal processing techniques like SNR optimization and advanced filtering algorithms highlighted by Freitas (2023). This approach minimises the influence of noise on the classification task at hand and ensures that the system can distinguish between genuine signals and interference. Also, with the introduction of GNNs, great promise has been shown in handling complex and noisy data, considering that they can leverage relationships between data points to achieve even higher classification accuracy than under interference (Khazane et al., 2024). Adaptations of ML algorithms that would make them ignore or compensate for environmental noise would significantly improve wireless network performance and reliability, especially in 5G environments.

Wireless ML systems depend on the crucial process of signal classification and modulation detection;

adversarial attacks pop up in every layer and cause misinterpretation of the signal (Mehrabi & Zhang, 2021). In this respect, robust modulation detection systems could be employed using techniques such as autoencoders and residual neural networks, among others, to enrich such models. For example, autoencoders learn a compressed representation of the data from the signal, which enables the model to "filter out" non-standard input that may represent adversarial interference (Reis, 2024). In contrast, residual networks can sustain performance under varying signal strengths and interference levels and, therefore, would be suitable for modulation detection in wireless networks. Using such techniques at the physical layer of wireless ML systems can increase signal classiﬁcation accuracy by a large gap. It might serve as a defence against malicious manipulation of signals.

Conventional anomaly detection mechanisms may fail to precisely detect adversarial activities in

real-time when the surroundings are very dynamic. The development of adaptive anomaly detection systems in real-time will provide immediate alerts and responses to threats, even when network conditions fluctuate (Wu et al., 2020). This can be done by incorporating reinforcement learning techniques that adaptively adjust the detection threshold referencing time-varying noise levels and network traffic patterns. The resulting systems would be more sensitive to the anomalies of wireless communication environments and thus enable ML models to adapt to preserve high accuracy under varying conditions, including those involving unexpected interference or jamming attacks.

Since ML model performance can degrade over time due to the constant evolution of threats, continual learning frameworks should be implemented where models keep updating themselves with the current state. In this approach, ML models gradually learn from new input data rather than depending only on a one-time training phase, thus enabling the system to recognise new attack patterns and environmental changes and respond quickly (Macas et al., 2024). While traditional model updates may involve significant system downtime, continual learning makes the update process progressive; during an update, the system remains up, and functionality is preserved (Usama et al., 2020). Applying this to wireless systems could mean the ML models remain resilient and adaptive to new vulnerabilities in response to discoveries of unanticipated data poisoning methods or innovative signal interference strategies.

## 

## 

## **Conclusion**

Integrated with wireless systems, ML improves signal classification and modulation,

improving network efficiency. This continuous improvement equally exposes wireless systems to an evolving security vulnerability, adversarial attacks and data poisoning that threatens the compromise of data's performance, reliability, and integrity (Chen et al., 2021; Soori et al., 2023). Such security threats are accentuated in this review, indicating severe flaws in the current defence mechanisms and stressing that more resilient and adaptive solutions must be developed.

While adversarial attacks – both noise manipulations and black-box attacks – continue to evolve,

adaptive defence strategies using advanced adversarial training and machine learning algorithms resilient to noise appear promising ways to make ML models resistant to interference from environmental inputs or with malicious data. Robust data validation techniques could be introduced into the architecture of federated learning to reduce the risks of data poisoning. At the same time, differential privacy is considered to help secure data integrity across distributed networks.

These findings indicate that anomaly detection and frameworks for real-time updating of ML models

have become even more essential in keeping up with dynamic environments. These will make it possible for wireless systems to grant security without a noticeable impact on system performance due to rapid adaptation against emerging threats (Wu et al., 2020; Macas et al., 2024). Therefore, this work should be oriented toward developing noise-resilient, adaptive defences able to maintain high accuracy and reliability under hostile conditions, especially in 5G and IoT applications.

Addressing such vulnerabilities with more robust ML models, combined with access to advanced security

techniques, will finally enable us to develop wireless systems that are efficient and resistant to continuously evolving cyber threats.

# 

# Appendix

## A1. Contributions table

| Name | Role | Task |
| --- | --- | --- |
| Noah | Project Leader | Literature review: topic 2 data poisoning  Research report: method, findings, code  Other: structure |
| Jemii | Co-Research Analyst | Literature review: introduction, topic 3 ML signal classification, gaps in research, conclusion  Research report: abstract, discussion  Other: proof-reading, project structure, alphabetical order references |
| Megan | Quality Assessor | Literature Review: Topic 1, Quality Assessor  Research Report: Executive Summary, Code Refactoring, Quality Assessor  Other: Progress Report, Minute Taker, Project Scope |
| Aleksyah | Co-Research Analyst | Research report: recommendations and conclusion |

## 

## A2. Log Book

**Weeks 1-3**

| Date | Week | Activity | Member | Status | Extra Comments |
| --- | --- | --- | --- | --- | --- |
| 6/8 | 1 | Choose Topic | All members | In Progress | Navigation app with improved services to what is currently available |
| 7/8 | 1 | Updated Meeting Minutes entries | Megan | Completed |  |
| 8/8 | 1 | Finalised Topic & Assembled Team | All members | Completed | We now have a group of 4 |
| 10/8 | 2 | Finding project Supervisor | All members | In Progress | Our supervisor is Kai |
| 14/8 | 2 | Updated Meeting Minutes entries | Megan | Completed |  |
| 15/8 | 2 | Changed the Project topic and Found the Project Supervisor | All members | Completed | The topic was changed to vulnerabilities in ML-integrated wireless systems |
| 16/8 | 2 | Draft of Project Requirements and Specifications section for Assignment 2 | Jemii | In Progress | I started to understand what was required to complete this project on-time |
| 16/8 | 2 | Draft of Project Plan | Megan | In Progress |  |
| 17/8 | 2 | Wrote introduction of Assignment 2 | Noah | In progress |  |
| 18/8 | 2 | Completed Time and Task Management for Assignment 2 | Megan | Completed |  |
| 18/8 | 2 | Completed Project Plan | Megan | Completed |  |
| 18/8 | 2 | Finalised Project Requirements and Specifications section | Jemii | Completed | Finished the section and ensured the rest of the group was happy with the quality |
| 18/8 | 2 | Draft of Professional Judgement and Decision-making started | Aleksyah | In Progress |  |
| 19/8 | 3 | Worked introduction again | Noah | Completed |  |
| 19/8 | 3 | Finished Project Plan for Assignment 2 | Megan | Completed |  |
| 20/8 | 3 | Worked on Draft of Professional Judgement and Decision-making |  |  |  |
| 20/8 | 3 | Draft conclusion for Assignment 2 | Jemii | In Progress | Based on what has been written, I wrote the conclusion |
| 21/8 | 3 | Updated Meeting Minutes entries | Megan | Completed |  |
| 22/8 | 3 | Updated Meeting Minutes entries | Megan | Completed | This was a meeting with the supervisor |
| 22/8 | 3 | Wrote about member roles and responsibilities | Noah | Completed |  |
| 23/8 | 3 | Finished Professional Judgement and Decision-making section | Aleksyah | Completed |  |
| 23/8 | 3 | Finished conclusion based on finished Assignment 2 | Jemii | Completed | Once everyone finished the project, I was able to finish the conclusion |
| 23/8 | 3 | Updated Assignment 2 Format | Megan | Completed |  |

**Weeks 4-7**

| Date | Week | Activity | Member | Status | Extra Comments |
| --- | --- | --- | --- | --- | --- |
| 29/8 | 4 | Started introduction for literature review | Jemii | In Progress | Started to get a general understanding of what the project should be discussed based on the research question |
| 30/8 | 4 | Finished draft introduction for literature review | Jemii | In Progress | Upon further research, I was able to complete the introduction |
| 3/9 | 5 | Draft objectives for literature review | Jemii | In Progress | Based on what was discussed in the team meeting, objectives could now be written |
| 3/9 | 5 | Began draft of the Objectives and Scope for the literature review | Megan | In Progress |  |
| 4/9 | 5 | Updated Meeting Minutes entries | Megan | Completed |  |
| 4/9 | 5 | Write comparative analysis of vulnerabilities occurring in various papers | Noah | In Progress |  |
| 5/9 | 5 | Researched mitigation strategies and recommendations | Aleksyah | In Progress |  |
| 6/9 | 5 | Draft of mitigation strategies section | Aleksyah | In Progress |  |
| 6/9 | 5 | Updated Meeting Minutes entries | Megan | Completed | This was a meeting with the supervisor |
| 6/9 | 5 | Fixed objectives based on feedback from the supervisor. | Jemii | Completed | Objectives and scope have now been combined. |
| 6/9 | 5 | Finalised Objectives and Scope based on supervisor feedback | Megan | Completed |  |
| 6/9 | 5 | Draft of Executive Summary | Megan | Completed |  |
| 6/9 | 5 | Begin Developing MATLAB Wireless Vulnerability Demonstration | Noah | In Progress | Using a simple feedforward network |
| 7/9 | 5 | Completed Executive Summary | Megan | Completed |  |
| 8/9 | 5 | Worked on mitigation strategies section | Aleksyah | In Progress |  |
| 9/9 | 6 | Turned executive summary into an abstract | Jemii | Completed | Literature reviews do not have an executive summary, so I changed it to an abstract |
| 10/9 | 6 | Research into data poisoning effects on wireless systems | Noah | In Progress |  |
| 11/9 | 6 | The introduction was fixed | Jemii | In Progress | Based on feedback from the supervisor last week, I could edit what was written. |
| 11/9 | 6 | Draft progress report for the mid-project review | Megan | In Progress |  |
| 11/9 | 6 | Draft 1 of mitigation strategies section | Aleksyah | In Progress |  |
| 13/9 | 6 | Finished comparative analysis of peer reviewed papers | Noah | Completed | Wrote comparative analysis for three papers however did not align with literature review structure |
| 17/9 | 7 | Updated progress report. I fixed up my contribution in the contributions table | Jemii | Completed | With all these changes made after the supervisor reviewed our work, I had to change my contribution to the table. |
| 17/9 | 7 | Completed progress report | Megan | Completed |  |
| 18/9 | 7 | Updated Meeting Minutes entries | Megan | Completed |  |
| 18/9 | 7 | Edited current draft | Aleksyah | In Progress | Feedback received from group on mitigations strategies |
| 19/9 | 7 | Mitigation strategies and recommendations fixed | Aleksyah | Completed | Section completed and finalised |
| 20/9 | 7 | Updated Meeting Minutes entries | Megan | Completed | This was a meeting with the supervisor. Supervisor stated that the project needed to be reworked into a research report, not a literature review |
| 20/9 | 7 | I fixed all my sections | Jemii | Completed | Based on the feedback in the meeting today, I changed things to fit more of a literature review structure |

**Weeks 8-12**

| Date | Week | Activity | Member | Status | Extra Comments |
| --- | --- | --- | --- | --- | --- |
| 23/9 | 8 | Updated Meeting Minutes entries | Megan | Completed |  |
| 24/9 | 8 | Began literary research on adversarial attacks | Megan | In Progress |  |
| 25/9 | 8 | Continue working on data poisoning topic | Noah | In Progress |  |
| 26/9 | 8 | Draft of Machine learning vulnerabilities in signal classification for the literature review | Jemii | In Progress | Fixed my section of the literature review |
| 26/9 | 8 | Began research on recommendations for the report | Aleksyah | In Progress |  |
| 27/9 | 8 | Completed literary research on adversarial attacks | Megan | Completed |  |
| 27/8 | 8 | Draft of comparative analysis on adversarial attacks | Megan | In Progress |  |
| 27/9 | 8 | Finalised Machine learning vulnerabilities in the signal classification section for the literature review | Jemii | Completed | I finished editing the section |
| 28/9 | 8 | Constructing MatLab Data poisoning demonstration | Noah | In Progress | SVM has better results then Feedforward neural network |
| 29/9 | 8 | Rough draft of recommendations | Aleksyah | In Progress |  |
| 3/10 | Stuvac | Working on Matlab Demo Anomaly detection module | Noah | In Progress |  |
| 3/10 | Stuvac | Completed comparative analysis on adversarial attacks | Megan | Completed |  |
| 4/10 | Stuvac | Updated Meeting Minutes entries | Megan | Completed |  |
| 11/10 | 9 | Updated Meeting Minutes entries | Megan | Completed | This was a meeting with the supervisor |
| 11/10 | 9 | Changed the structure of the literature review | Jemii | Completed | Changed the whole structure of the literature review. I put in new titles. |
| 12/10 | 9 | Wrote Gaps In Research, introduction, and conclusion for the literature review | Jemii | In Progress | I added information to the new structure of the literature review. I was in charge of fixing the whole review based on feedback from the supervisor and research into what a structure of it looks like |
| 14/10 | 10 | Working on Matlab demonstration data poisoning section | Noah | Completed | Adding performance matrix later |
| 16/10 | 10 | Updated Meeting Minutes entries | Megan | Completed |  |
| 16/10 | 10 | Finished the draft of the literature review and proof-read it before sending it to our supervisor for feedback | Jemii | Completed | Completed the literature review |
| 16/10 | 10 | Revised written sections | Megan | Completed |  |
| 16/10 | 10 | Draft 1 of recommendations | Aleksyah | In Progress |  |
| 17/10 | 10 | Started on the research report. | Jemii | In Progress | I am doing the discussion section. I looked over the practical part of the report to see how I would tackle this section. I wrote notes of what I should be talking about. |
| 18/10 | 10 | Wrote up a rough draft for the discussion section | Jemii | In Progress | Started on the discussion section of the research report |
| 19/10 | 10 | Draft of executive summary | Megan | In Progress |  |
| 19/10 | 10 | Finished the discussion section | Jemii | Completed | I finished as much as I could for the discussion section |
| 19/10 | 10 | Read through the literature review | Aleksyah | In Progress | Added relevant information to recommendations for research report to tie it together |
| 20/10 | 10 | Introduced Performance Metrics and error checking techniques to evaluate the performance of the machine learning model | Noah | In Progress | Added F1 Score and Checksum values |
| 22/10 | 11 | Wrote draft 2 of recommendations for research report | Aleksyah | In Progress |  |
| 23/10 | 11 | Refactored Matlab Code | Megan | Completed | Made minor changes to the code |
| 24/10 | 11 | Work on method in research report | Noah | In Progress |  |
| 24/10 | 11 | Rewrote the abstract | Jemii | In Progress | Based on our new structure and discussion, I changed the abstract to fit the report better. |
| 25/10 | 11 | Revised written sections | Megan | Completed | Re-read all sections I wrote to make sure I was happy with the content and structure |
| 27/10 | 11 | Fixed challenges | Jemii | In Progress | Based on what was written in the findings section, I added and changed part of the challenges section. |
| 28/10 | 12 | Implemented both SVM and NN | Noah | In Progress | Couldn't fit the dataset with RNN network so used SVM and nn to display traditional and deep learning methods |
| 29/10 | 12 | Finished recommendations for research report | Aleksyah | Completed |  |
| 29/10 | 12 | I re-read all my sections to ensure I was happy with it | Jemii | In Progress | Before my group members read over my work, I ensured I was happy with what I had written |
| 30/10 | 12 | Read the whole report and began the conclusion for the research report | Aleksyah | In -Progress |  |
| 31/10 | 12 | Finished the conclusion for the research report | Aleksyah | Completed |  |
| 31/10 | 12 | Completed MatLab demonstration for SVM and NN | Noah | Completed | Added more graphs and more functions such as grapher, plotSig, checksum |
| 31/10 | 12 | Completed the Executive Summary | Megan | Completed |  |
| 31/10 | 12 | Read over the report as a whole | Jemii | Completed | I ensured I was happy with the overall quality of the literature review and research report. |
| 31/10 | 12 | Proofread and Finalised Report for Submission | Megan | Completed |  |

# References

Adesina, D., Hsieh, C. C., Sagduyu, Y. E., & Qian, L. (2022). Adversarial machine learning in wireless communications using RF data: A review. IEEE Communications Surveys & Tutorials, 25(1), 77–100. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9887796

Akter, S., Dwivedi, Y. K., Sajib, S., Biswas, K., Bandara, R. J., & Michael, K. (2022). Algorithmic bias in machine learning-based marketing models. Journal of Business Research, 144, 201–216. https://doi.org/10.1016/j.jbusres.2022.01.083

Ali, S., Saad, W., & Steinbach, D. (2020). White paper on machine learning in 6G wireless communication networks. 6G Research Visions. http://urn.fi/urn:isbn:9789526226736

Alhajjar, E., Maxwell, P., & Bastian, N. (2021). Adversarial machine learning in network intrusion detection systems. Expert Systems with Applications, 186. https://doi.org/10.1016/j.eswa.2021.115782

Apnic Foundation. (2022). Adversarial machine learning attacks in wireless networks. https://apnic.foundation/projects/adversarial-machine-learning-attacks-in-wireless-networks/

Akyildiz, I. F., Lee, W.-Y., & Chowdhury, K. R. (2009). CRAHNs: Cognitive radio ad hoc networks. Ad Hoc Networks, 7(5), 810–836. https://doi.org/10.1016/j.adhoc.2009.01.001

Aslan, Ö., Aktuğ, S. S., Ozkan-Okay, M., Yilmaz, A. A., & Akin, E. (2023). A comprehensive review of cybersecurity vulnerabilities, threats, attacks, and solutions. Electronics, 12(6), 1–42. https://www.mdpi.com/2079-9292/12/6/1333

Bagdasaryan, E., Veit, A., Hua, Y., Estrin, D., & Shmatikov, V. (2020). How to backdoor federated learning. Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics (AISTATS).

Bhagoji, A. N., Chakraborty, S., Mittal, P., & Calo, S. B. (2019). Analyzing federated learning through an adversarial lens. International Conference on Machine Learning (ICML).

Boesch, G. (2023, December 2). What is adversarial machine learning? Attack methods in 2024. viso.ai. https://viso.ai/deep-learning/adversarial-machine-learning/#:~:text=An%20adversarial%20attack%20is%20a,valid%20input%20to%20a%20human

Bonneau, J., Herley, C., Van Oorschot, P. C., & Stajano, F. (2012). The quest to replace passwords: A framework for comparative evaluation of web authentication schemes. In IEEE Symposium on Security and Privacy. https://doi.org/10.1109/SP.2012.44

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys (CSUR), 41(3), 1–58. https://doi.org/10.1145/1541880.1541882

Chen, Z., Cui, H., Xiang, J., Qiu, K., Huang, L., Zheng, S., Chen, S., Xuan, Q., & Yang, X. (2021). SigNet: A novel deep learning framework for radio signal classification. IEEE Transactions on Cognitive Communications and Networking, 8(2), 529–541. https://doi.org/10.1109/tccn.2021.3120997

Cosma, G., Brown, D., Archer, M., Khan, M., & Graham Pockley, A. (2017). A survey on computational intelligence approaches for predictive modeling in prostate cancer. Expert Systems with Applications, 70, 1–19. https://doi.org/10.1016/j.eswa.2016.11.006

Dhiraj, K. (2019, November 21). Top 4 advantages and disadvantages of support vector machine or SVM. Medium. https://dhirajkumarblog.medium.com/top-4-advantages-and-disadvantages-of-support-vector-machine-or-svm-a3c06a2b107

Elbir, A. M., & Papazafeiropoulos, A. (2020). Deep learning for channel estimation: A comprehensive survey. IEEE Communications Surveys & Tutorials, 22(3), 1985–2015. https://doi.org/10.1109/COMST.2020.2992123

Freitas, P. (2023, April 19). Threats to machine learning-based systems – Part 2 of 5. Side Channel – Tempest. https://www.sidechannel.blog/en/threats-to-machine-learning-based-systems-part-2-of-5/

Goldblum, M., Fowl, L., & Goldstein, T. (2020). Adversarial attacks on machine learning systems: Practical implications for cybersecurity. Journal of Cybersecurity.

Goodfellow, I. J., Shlens, J., & Szegedy, C. (2014). Explaining and harnessing adversarial examples. arXiv preprint arXiv:1412.6572. https://doi.org/10.48550/arXiv.1412.6572

Haykin, S., & Widrow, B. (Eds.). (2002). Least-mean-square adaptive filters. Wiley. https://doi.org/10.1002/0471224229

Hebbar, P. (2023, May 29). Attacking machine learning with adversarial examples. Medium. https://medium.com/@Pran-Ker/attacking-machine-learning-with-adversarial-examples-ef883e3273c5

Hodo, E., Bellekens, X., Hamilton, A., Tachtatzis, C., & Atkinson, R. (2016). Threat analysis of IoT networks using artificial neural network intrusion detection system. In 2016 International Symposium on Networks, Computers and Communications (ISNCC) (pp. 1–6). IEEE. https://doi.org/10.1109/ISNCC.2016.7746067

Huang, X., Kwiatkowska, M., Wang, S., & Wu, M. (2017). Safety verification of deep neural networks. In International Conference on Computer Aided Verification. Springer, Cham. https://doi.org/10.1007/978-3-319-63387-9\_5

Hulayyil, S. B., Li, S., & Xu, L. (2023). Machine-learning-based vulnerability detection and classification in internet of things device security. Electronics, 12(18), 3927. https://doi.org/10.3390/electronics12183927

Ibitoye, O., Abou-Khamis, R., Shehaby, M. E., Matrawy, A., & Shafiq, M. O. (2019). The threat of adversarial attacks on machine learning in network security: A survey. arXiv. https://doi.org/10.48550/arXiv.1911.02621

Ifran, M. M., Ali, S., Yaqoob, I., & Zafar, N. (2021). Towards deep learning: A review on adversarial attacks. 2021 International Conference on Artificial Intelligence (ICAI), 91–96. https://ieeexplore.ieee.org/abstract/document/9445247#full-text-header

Jagielski, M., Oprea, A., Biggio, B., Liu, C., Nita-Rotaru, C., & Li, B. (2018). Manipulating machine learning: Poisoning attacks and countermeasures. Proceedings of the IEEE Security and Privacy Workshops.

Kanade, V. (2022, August 30). What is machine learning? Definition, types, applications, and trends for 2022. Spiceworks. https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-ml/

Katz, G., Barrett, C., Dill, D. L., Julian, K., & Kochenderfer, M. J. (2017). Reluplex: An efficient SMT solver for verifying deep neural networks. In Computer Aided Verification (CAV), 97–117. https://doi.org/10.1007/978-3-319-63387-9\_5

Khazane, H., Ridouani, M., Salahdine, F., & Kaabouch, N. (2024). A holistic review of machine learning adversarial attacks in IoT networks. Future Internet, 16(1), 32. https://doi.org/10.3390/fi16010032

Kumar, N., Adhikari, S., Sharma, M., & Bawa, M. (2023). Attacks on artificial intelligence in the healthcare industry: Challenges and solutions. Computational Intelligence and Neuroscience, 2023, Article 1–15. https://doi.org/10.1155/2023/2431256

Lee, M., Park, J., & Chang, J. (2022). Enhancing security in Internet of Things networks using adversarial machine learning techniques. Sensors, 22(8), 3121. https://doi.org/10.3390/s22083121

Li, X., Cheng, J., & Zhang, T. (2020). A survey on adversarial attacks and defenses in deep learning. IEEE Transactions on Neural Networks and Learning Systems, 31(12), 4092–4115. https://doi.org/10.1109/TNNLS.2020.2991772

Lin, M., Ma, L., & Liu, H. (2021). Wireless network security: A review of machine learning methods. Computer Networks, 196, 108224. https://doi.org/10.1016/j.comnet.2021.108224

Liu, Y., Yang, C., Li, Y., & Wang, T. (2022). Deep learning-based radio signal recognition: A comprehensive survey. IEEE Access, 10, 11740–11757. https://doi.org/10.1109/ACCESS.2022.3144497

Lopes, P., & Ponte, L. (2020). Federated learning: Adversarial attacks and defenses. IEEE Transactions on Mobile Computing, 1–15. https://doi.org/10.1109/TMC.2020.3021701

Mishra, V., & Shukla, A. (2021). Mitigating adversarial attacks on deep learning models. Pattern Recognition Letters, 150, 150–157. https://doi.org/10.1016/j.patrec.2021.07.017

Mittal, S., Sharma, A., & Jain, A. (2023). Adversarial attacks in machine learning: A comprehensive review. Journal of Artificial Intelligence Research, 76, 471–511. https://doi.org/10.1613/jair.1.12129

Mowery, K., & Shacham, H. (2023). A survey of threats to machine learning systems. IEEE Security & Privacy Magazine, 21(2), 68–80. https://doi.org/10.1109/MSP.2023.3234766

Neto, G., Cardoso, J., & Costa, P. (2023). Improving security in smart city applications with adversarial training. Smart Cities, 6(1), 88–105. https://doi.org/10.3390/smartcities6010008

Papernot, N., McDaniel, P., Goodfellow, I., Jha, S., Celik, Z. B., & Swami, A. (2016). Practical black-box attacks against machine learning. Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security (AsiaCCS), 506–519. https://doi.org/10.1145/3052973.3053009

Qi, X., & Yang, X. (2022). Adversarial machine learning in natural language processing: A survey. IEEE Access, 10, 27155–27170. https://doi.org/10.1109/ACCESS.2022.3155725

Qiu, W., & Tian, Y. (2020). Review of machine learning for wireless communications. IEEE Communications Surveys & Tutorials, 22(4), 2224–2251. https://doi.org/10.1109/COMST.2020.3003108

Rahman, M., & Islam, S. M. S. (2022). Adversarial attacks on deep learning for computer vision: A survey of current challenges and future research directions. Sensors, 22(9), 3102. https://doi.org/10.3390/s22093102

Rana, S., Gupta, M., Vyas, P., & Vyas, A. (2021). Security in smart grid applications: A review of adversarial attacks and defenses. Energy Reports, 7, 144–152. https://doi.org/10.1016/j.egyr.2020.12.039

Saha, M., & Banerjee, S. (2023). A review on adversarial machine learning: Recent advancements and future directions. International Journal of Machine Learning and Cybernetics, 14(1), 15–29. https://doi.org/10.1007/s13042-022-01510-7

Sharma, D., & Singh, R. (2021). AI in healthcare and cybersecurity: Reviewing the use of machine learning models in addressing security threats. Healthcare Informatics Research, 27(2), 130–140. https://doi.org/10.4258/hir.2021.27.2.130

Su, J., Vargas, D. V., & Sakurai, K. (2019). One pixel attack for fooling deep neural networks. IEEE Transactions on Evolutionary Computation, 23(5), 828–841. https://doi.org/10.1109/TEVC.2019.2890851

Sun, J., Liu, Y., & Zhang, H. (2021). Research progress in machine learning for wireless communication. Mobile Networks and Applications, 26(1), 53–66. https://doi.org/10.1007/s11036-020-01635-8

Thakur, P., & Dogra, D. P. (2020). Security challenges in IoT and adversarial machine learning: A survey. Internet of Things, 12, 100270. https://doi.org/10.1016/j.iot.2020.100270

Vakharia, V., Shah, A., & Doshi, N. (2023). Adversarial attacks in machine learning and countermeasures: A review. Machine Learning and Knowledge Extraction, 5(1), 12–25. https://doi.org/10.3390/make5010002

Wang, Q., & He, J. (2022). Mitigating adversarial attacks on deep learning models using randomized smoothing. Journal of Artificial Intelligence Research, 75, 345–365. https://doi.org/10.1613/jair.1.13057

Xu, W., Evans, D., & Qi, Y. (2016). Automatically evading classifiers: A case study on PDF malware classifiers. In Network and Distributed Systems Security (NDSS) Symposium 2016.

Yan, H., & Ding, H. (2023). Federated learning with adversarial attacks: A review. Journal of Information Security and Applications, 65, 103187. https://doi.org/10.1016/j.jisa.2022.103187

Yang, Y., & Wu, D. (2022). A survey on adversarial attacks and defenses for machine learning models in computer vision. IEEE Access, 10, 102227–102241. https://doi.org/10.1109/ACCESS.2022.3193157

Zhang, Y., & Yu, H. (2020). Adversarial machine learning: A literature survey. arXiv preprint arXiv:2007.07392. https://doi.org/10.48550/arXiv.2007.07392