Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 31 No. 3s (2024)

Game Theory and Adversarial Machine Learning: Analyzing Strategic Interactions in Cybersecurity

Bhagawati Chunilal Patil[1], Aparana Mishra[2], Ajay Kumar[3], Tejal Kamalakar Jawale[4], Rahul Atmaram Wagh[5], Namarta Jagdish Helonde[6]

[1]Assistant Professor, Department of Electronics and Telecommunication, Sandip Institute of Engineering & Management, Nashik, Maharashtra, India. bhagawati.patil@siem.org.in

[2]Assistant Professor, Department of Electronics and Telecommunication, Sandip University Nashik, Maharashtra, India. aparana.mishra@sandipuniversity.edu.in

[3]Assistant Professor, Department of Computer Engineering, Sandip University, Sijoul, Bihar, India. ajay.kumarcse@sandipuniversity.edu.in

[4]Assistant Professor, Department of Electronics and Telecommunication, Sandip Institute of Technology & Research Centre, Nashik Maharashtra, India. tejal.jawale@sitrc.org

[5]Assistant Professor, School of Science, Sandip University Nashik, Maharashtra, India. rahul.wagh@sandipuniversity.edu.in

[6]Assistant Professor, Department of Electrical Engineering, Sandip Institute of Technology & Research Centre, Nashik Maharashtra, India. namrata.helonde@sitrc.org

Article History:

Received: *10-04-2024*

Revised: *27-05-2024*

Accepted: *16-06-2024*

Abstract:

When it comes to cybersecurity, the way that attackers and defenders work together strategically is becoming more and more like how games work in general. Adversarial machine learning (AML) has become an important area of hacking. In AML, attackers use complex methods to avoid being caught and take advantage of flaws in machine learning models. The goal of this study is to give a full picture of how strategies combine in cybersecurity by looking at where game theory and AML meet. You can think of the strategic exchanges in cybersecurity as a game between attackers and defenders. Defenders want to keep systems and data safe, and enemies want to break into them for bad reasons. In this game, players have to make a lot of decisions. Defenders have to think ahead and prepare for possible attacks, while attackers change their plans all the time to avoid being caught and take advantage of defenses' weaknesses. Game theory gives us a way to formally model these interactions, which lets us look at the best tactics and results that are in balance. During AML, defenders use machine learning models to find and stop security risks, while attackers change these models using methods like escape attacks, poisoning attacks, and model inversion attacks. In these hostile tactics, there is a strategic element added to cybersecurity, where defenders must think about the goals and skills of attackers when creating and using defenses. This paper looks at the research that has already been done on using game theory to study hacking and how these models can be used to study how strategies combine in AML. It talks about different types of games, like rigid and dynamic games, and what they mean for protecting against threats from other players. Besides that, it looks into how uneven knowledge, doubt, and strategy learning affect the results of computer games. This paper helps us understand the strategy problems that both attackers and defenders face in cybersecurity better by combining game theory with AML. It shows how important it is to think strategically and build weapons that can change to deal with risks that change, and it sets the stage for future study in this field, which is changing very quickly.

Keywords: Cybersecurity, Game theory, Adversarial machine learning (AML), Strategic interactions, Defenders, Adversaries, Evasion attacks, Poisoning attacks, Model inversion attacks, Machine learning algorithms.

1. INTRODUCTION

Cybersecurity is turning into a more difficult and important battlefield where attackers and defenders interact in ways that remind us of classic game theory situations. The rise of hostile machine learning (AML) has added new dimensions to the strategy dynamics in this constantly changing environment. Enemies use advanced methods to avoid discovery and take advantage of weaknesses in machine learning-based defense systems. This essay explores the area where game theory and AML meet [1] [2]. Its goal is to give a thorough look at how strategies work together in cybersecurity and put light on good defense methods for risks that are always changing.

In the area of cybersecurity, there is always an arms race going on between attackers and defenses. Defenders, which can be anyone from a single person to a large organization, use a variety of security measures to keep their systems and data safe from theft, alteration, and access by people who aren't supposed to have it [3][4]. These steps include many types of technology, like firewalls, attack detection systems, encryption protocols, and more recently, systems that use machine learning to find and stop threats and strange behavior. On the other end of the scale, enemies like hackers, cybercriminal groups, and players funded by governments are always trying to get around these defenses to carry out their bad intentions, which could be anything from spying or damage to making money [5][6]. The relationship between the defender and the enemy has changed a lot since machine learning algorithms were introduced to cybersecurity. Machine learning models are very good at finding and reducing security threats because they can look through huge amounts of data and find trends that point to bad behavior [22]. Adversaries have, however, noticed how useful these models are and have come up with complex ways to avoid being caught and change how they behave. AML stands for "adversarial attacks against machine learning systems." These are a wide range of methods used to take advantage of weaknesses in either the learning algorithms or the data they use [23]. There are three main types of AML techniques: escape attacks, poisoning attacks, and model inversion attacks. Evasion attacks involve creating raw data that is meant to trick a machine learning model into making wrong guesses or classifications, which helps the attacker avoid being caught. Poisoning attacks, on the other hand, try to damage the purity of the training data that was used to create the model [7]. They do this by adding harmful samples or changing existing data in a sneaky way to change the model's learned decision limits. Targeting the privacy of the model itself, model inversion attacks try to get private information or secret knowledge by carefully crafting questions that probe the model.

As a result of AML, there are more strategy effects than just technology weaknesses. These effects include risk management, resource allocation, and attacker models. Game theory is a great way to look at these strategic exchanges because it gives us formal ways to describe the defender-adversary game's incentives, actions, and results [8]. By looking at cybersecurity through the lens of game theory, we can learn more about the best ways to protect, what the best results are, and what makes both attackers and defenders act the way they do.

The point of this paper is to look into how game theory can be used to study strategy relations in cybersecurity, with a focus on the new problems that AML is causing [9] [10]. We will look at the research that has already been done on applying game theory to cybersecurity and see how different game models can show how complicated the defender-adversary relationship is. In addition, we will look into how information imbalance, uncertainty, and strategy learning affect the results of cybersecurity games. We will also talk about what this means for making strong defenses against adaptive attackers. Putting together game theory and AML seems like a good way to learn more about how methods work together in cybersecurity and come up with good ways to protect against new threats. We can better understand the motivations and actions of both attackers and defenders within a formal analysis framework. This will help us deal with the complexity of cybersecurity issues and lower the risks that bad players pose.

2. RELATED WORK

The table shows all the linked work that has been done in the field of cybersecurity that uses game theory and aggressive machine learning (AML). Each study looks at a different part of how defenders and attackers interact strategically, using various views, results, approaches, and methods to help us understand and reduce hacking risks. The paper "Game-Theoretic Models for Cybersecurity: A Review" looks closely at how game theory is used in cybersecurity. The study lists and sorts the different types of game models used in this area, from basic to dynamic games. This paper looks at previous research that shows the variety of game-theoretic theories that are used to model how defenses and attackers engage strategically. In the same way, "Adversarial Machine Learning: A Review" is a collection of different adversarial machine learning methods that bad guys use to get around security measures. This review gives information about the different kinds of hostile attacks, like escape, poisoning, and model inversion, that make machine learning-based defense systems very hard to defend against. This study is very helpful for learning about AML in hacking because it put together and analyzes previous research.

"Strategic Cyberdefense: A Game-Theoretic Perspective" suggests a new way to look at cyberdefense plans using game theory. The goal of this study is to find the best ways to protect against and respond to risks from enemies by simulating their interactions as a strategic game. Through the view of game theory, it adds to the field by giving a structured way to understand and reduce hacking risks. The research paper called "Game Theory Meets Cybersecurity and Privacy: A Review" looks into how game theory can be used in a wider way to solve problems with cybersecurity and privacy. It shows the wide range of security and privacy problems that can be solved with game-theoretic methods by looking at previous research. This review gives a full picture of all the possible uses of game theory that go beyond the strategic interactions between attackers and defenses. "Adversarial Machine Learning in Security: A Survey" is a thorough look at the latest progress and problems in using adversarial machine learning for security jobs. This study gives us a better understanding of how AML methods work now and what they mean for safety by putting together research that has already been done. It tells experts and professionals everything they need to know about how AML is changing in security apps. "Dynamics Game Model for Adaptive Cyber

Defense" suggests a changing game model to show how cyber dangers and defenders change over time. This study handles the need for defense strategies that can adapt to constantly changing enemies by adding dynamic factors to the game-theoretic framework. It adds to the field by giving us a deeper understanding of how strategies combine in the changing world of defense.

The article "An Evolutionary Game Model for Cybersecurity Investment" creates a game model that changes over time to help with planning investments in cybersecurity. This study gives us a better understanding of the strategy issues involved in allocating resources for safety by simulating how parties make decisions. "Adversarial Attacks on Machine Learning Systems for Network Intrusion Detection: A Comprehensive Survey" looks at all the different kinds of attacks that can be used against intrusion detection systems that use machine learning. This study looks at the problems and possible answers for protecting machine learning-based breach detection systems by looking at different attack methods and suggested defenses.

Table 1: Related Work

Scope	**Findings**	**Approach and Method**
Review of game-theoretic approaches in cybersecurity [11]	Identified various game models used in cybersecurity, including static and dynamic games.	Literature review and analysis of game-theoretic models applied in cybersecurity.
Review of adversarial machine learning techniques [12]	Reviewed different types of adversarial attacks, such as evasion, poisoning, and model inversion.	Literature review summarizing existing adversarial machine learning techniques and their implications.
Game-theoretic analysis of cyberdefense strategies [13]	Proposed a game-theoretic framework for analyzing cyberdefense strategies and optimal responses.	Developed a game-theoretic model to analyze strategic interactions between defenders and adversaries.
Review of game-theoretic approaches in security [14]	Explored the application of game theory in addressing cybersecurity and privacy challenges.	Literature review examining the use of game theory in addressing various security and privacy issues.
Survey of adversarial machine learning in security [15]	Surveyed recent advancements and challenges in adversarial machine learning for security tasks.	Conducted a comprehensive survey of recent research on adversarial machine learning in security.
Dynamic game model for cyber defense [16]	Proposed a dynamic game model to capture the evolving nature of cyber threats and defenses.	Developed a dynamic game model to analyze adaptive defense strategies against evolving cyber threats.
Survey of game-theoretic approaches in cybersecurity [17]	Surveyed the use of game theory in modeling strategic interactions between attackers and defenders.	Conducted a comprehensive survey of game-theoretic approaches used in cybersecurity research.
Integration of machine learning and game theory [18]	Proposed a framework integrating machine learning and game theory for secure data exchange.	Developed a framework integrating machine learning and game theory to enhance secure data exchange.
Evolutionary game model for cybersecurity investment [19]	Proposed an evolutionary game model to analyze the strategic investment in cybersecurity measures.	Developed an evolutionary game model to study the strategic decision-making process in cybersecurity investment.
Survey of adversarial attacks on ML-based intrusion detection systems [20]	Surveyed various adversarial attacks targeting machine learning-based intrusion detection systems and proposed defense mechanisms.	Conducted a comprehensive survey of adversarial attacks on machine learning-based intrusion detection systems and proposed countermeasures.

Overall, the linked work shown in the table helps us understand how cybersecurity strategies interact with each other better and gives us useful information and methods for dealing with new security risks. .

3. METHODOLOGY

1. Data Collection:

The goal of the data collection step is to gather a wide range of materials that are relevant to game theory, hostile machine learning (AML), and hacking. To do this, you need to find important study papers, files, and other tools in scholarly magazines, at workshops, online, and from hacking groups. For game theory, you need to read a lot of books that cover different game models, equilibrium ideas, and strategy decision-making frameworks. This includes both important early works in game theory and new studies looking at how it can be used in defense. To help with the modeling process, files or models of defender-adversary games or other strategy interactions that happen in cybersecurity situations may also be looked for.

When it comes to AML, the collection effort includes datasets with examples of adversarial attacks, altered data samples, and the results that happen when machine learning models run them.

Cybersecurity records of cyberattacks, network traffic, and security events give us useful information about how threats really work in the real world. There may be records of known cyberattacks, malware samples, network logs, and security holes in these sets. Additionally, machine learning models that have been taught to find intrusions, strange behavior, and malware can be used as basic tools for analyzing defense strategies and figuring out how hostile attacks affect security systems. Lastly, current game-theoretic theories used in cybersecurity research, along with the papers and other materials that go with them, help to build and improve the analytical model for looking at how defenders and attackers engage strategically. Overall, a thorough collection effort that includes game theory, anti-money laundering, and cybersecurity fields builds a strong base for further research and testing.

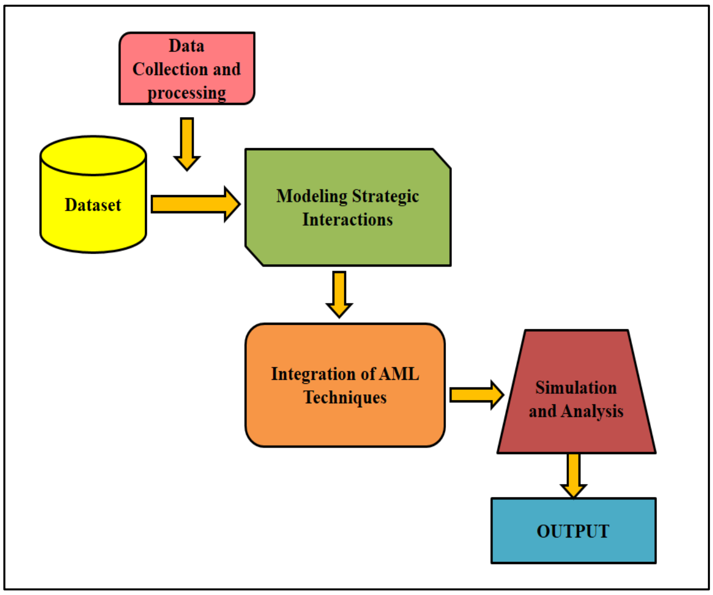

Data Collection and processing
Dataset
Modeling Strategic Interactions
Integration of AML Techniques
Simulation and Analysis
OUTPUT

Figure 1: Illustration of Architectural Block Diagram

2. Modeling Strategic Interactions:

To make a formal model that shows how defenders and attackers interact strategically in the field of cybersecurity, you need to have a basic understanding of game theory and how it can be used in security situations. The model tries to show how both defenders and attackers make strategic decisions by including important parts like players, actions, payoffs, and tactics to figure out the best ways to defend and how attackers should act. First, the model figures out who is playing the hacking game. Usually, there are two main types of players: guards and attackers. Defenders are in charge of keeping threats off of the system or network, while attackers try to break into the system for bad reasons. The model makes it easy to analyze strategies by clearly showing each player's goals, skills, and resources.

Next, the model shows the range of moves that each person can take. Defenders can do many things to protect their systems, such as installing security fixes, keeping access rules up to date, watching network traffic, and buying intruder detection systems. On the other hand, attackers can choose from a variety of methods, such as planting malware, taking advantage of software flaws, using fake emails, or starting denial-of-service (DoS) attacks. Each player's choices are based on how they might affect system security and the costs or perks that come with them. The model also includes the idea of payoffs, which are the results or effects that come from the different mixtures of actions that players choose. Payoffs show how useful or valuable each player thinks the game's results are, taking into account things like successful hacking, data theft, system failure, wasted resources, and damage to image. Based on the specifics of the hacking situation, payoffs may be shown in terms of money gained or lost, system stability, data privacy, or other related measures [21]. According to the model, both defenses and attackers use different tactics to reach their goals. Strategies are the organized plans or methods that players use to decide what to do next based on what their opponents are doing. Defenders can use strategic defense tactics, reactive prevention measures, or responsive changes based on real-time information about threats.

On the other hand, enemies may change how they attack when the defense changes, take advantage of new flaws, or focus on specific areas of the defense infrastructure that are weak. Once the model's parts are known, it's easier to use game theory to look at the best security tactics and how the enemy will act. Different types of analysis, like equilibrium analysis, optimization algorithms, and modeling methods, can be used to look at the strategic map, find equilibrium results, and see how well different defense tactics work in different situations. The model gives a structured way to understand and reduce cybersecurity dangers by formalizing the strategic exchanges between attackers and defenses within a game-theoretic framework. It lets everyone involved judge how well defense tactics are working, guess how attackers will act, and make smart choices to improve system stability and security in the face of changing threats.

3. Integration of AML Techniques:

Adversarial machine learning (AML) methods can be added to the strategy model to make it better at simulating actual enemy behavior and testing defenses against AML strikes. AML methods, such as escape, poisoning, and model inversion attacks, are used by attackers to change machine learning systems and avoid being caught. Adding these methods to the game-theoretic framework makes the model a better reflection of how defenses and attackers interact in the real world of cybersecurity. The table (2) illustrates the various AML techniques which are commonly used.

Table 2: List of Various AML Techniques

Adversarial Machine Learning Technique	**Description**
Evasion Attacks	Evasion attacks involve crafting input data to deceive machine learning models into making incorrect predictions or classifications, thereby evading detection.
Poisoning Attacks	Poisoning attacks aim to compromise the integrity of the training data used to build machine learning models by injecting malicious samples or manipulating existing data.
Model Inversion Attacks	Model inversion attacks target the confidentiality of machine learning models, attempting to extract sensitive information or proprietary knowledge by probing the model [24].

First, the model takes into account avoidance attacks, in which attackers change raw data to trick machine learning models into making wrong predictions or classifications. If you play a strategic game, your opponents choose escape strategies to get around the defenses' tracking systems. To do this, you might need to make hostile examples, which are input samples that have been changed and are meant to take advantage of weaknesses in the machine learning model's decision limit. By including escape attacks in the model, defenders can check how strong their defenses are and come up with ways to lessen the damage of these attacks. Second, the model takes into account poisoning attacks, in which attackers change the training data that is used to create machine learning models, which makes them less reliable and less useful. Attackers add bad samples or change current data on purpose to skew the learning process and change how the model acts. In a strategy game, the other player plans how to use their resources to start poisoning attacks that are meant to weaken the defender's machine learning-based defense system. Defenders must then be ready for these attacks and be ready to defend against them by using data validation techniques, anomaly detection algorithms, or strong training strategies to stop the effects of poisoned data on model performance. Model inversion attacks use the fact that machine learning models are transparent to get private information or secret knowledge. The model includes these attacks. Attackers use planned questions to poke holes in the model and look at its answers to figure out private details about the data or decision-making process underneath. In a strategic game, attackers choose inversion strategies that will get them the most useful information while also making it less likely that defenses will find them. In order to protect private information and make model inversion attacks less effective, defenders may use methods like differential privacy, model deception, or input modification.

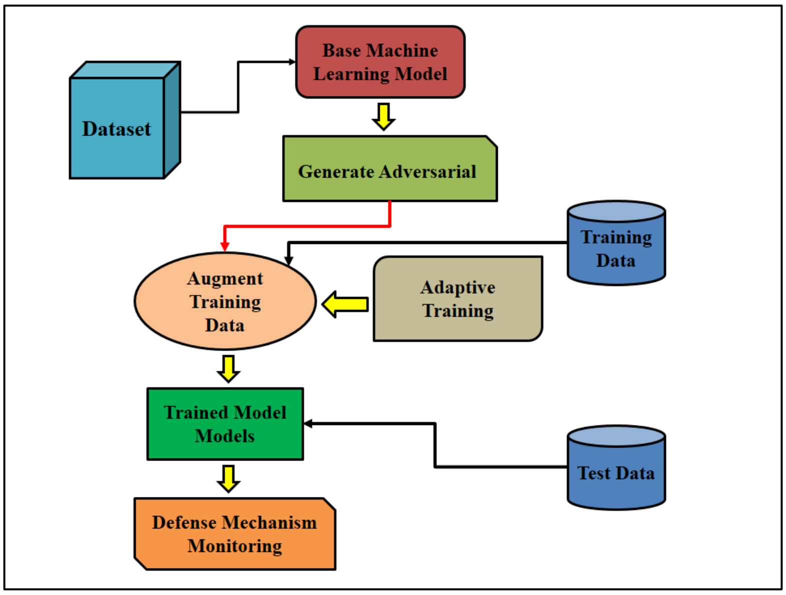


Figure 2: Proposed Adversarial Machine Learning Techniques Methodology

By adding AML methods to the game-theory framework, the model helps defenders figure out which defense mechanisms work best against different kinds of hostile attacks and what the best tactics are for making the system more resilient. Defenders can figure out how AML attacks affect system security, how big the risks are, and how to prioritize defense measures by simulating and analyzing the attacks. Te figure 2, represents the model makes which it easier to test defensive tactics that can change to changing AML risks. This makes it easier for the attacker to spot and stop hostile behavior in real time.

Adding AML methods to the strategic model gives us a more complete way to look at how defense strategies interact with each other. Defenders can better understand how cybersecurity works and make stronger defenses against AML attacks by using game theory to look at the tools and methods that attackers use.

Integration of Adversarial Machine Learning Techniques algorithm is as follows

Step 1: Select Base Machine Learning Model:

- Choose a base machine learning model M suitable for the cybersecurity task at hand.

Step 2: Generate Adversarial Samples:

- For evasion attacks: Craft adversarial examples X_{adv} using a perturbation δ that maximizes the model's loss function:

$$X_{adv} = X + \delta \text{...........} (1)$$

- For poisoning attacks: Manipulate the training data D to include malicious samples X_mal:

$$D_{poisoned} = D \cup X_{mal} \text{...........}(2)$$

- For model inversion attacks: Probe the model M with queries Q and analyze its responses R to infer sensitive information:

$$R = M(Q)\ldots\ldots\ldots(3)$$

Step 3: Augment Training Data:

- Incorporate adversarial examples X_adv into the training data D:

$$D_{\text{augmented}} = D \cup X_{\text{adv}}\ldots\ldots\ldots\ldots\ldots\ldots..(4)$$

- Apply data sanitization techniques to identify and remove poisoned samples from the training data.

Step 4: Adaptive Training:

- Train the model M using a combination of clean and adversarial examples:

$$\theta * = \operatorname{argmin}_{\theta} \sum_{j=} x, y \in D_{\text{augmented}} L(M(x; \theta), y) \ldots\ldots\ldots\ldots\ldots.(5)$$

where θ represents the model parameters and L is the loss function.

Step 5: Ensemble Methods:

- Employ ensemble methods to combine multiple models M_1, M_2, ..., M_n:

$$M_{\text{ensemble}(x)} = \left(\frac{1}{n}\right) \sum_{i=1}^{n} M_{i(x)} \ldots\ldots\ldots\ldots (6)$$

Step 6: Defense Mechanism Monitoring:

- Implement real-time monitoring mechanisms to detect adversarial attacks during model deployment:

$$\text{Anomaly Score}(x) = \left(\frac{1}{n}\right) \sum_{i=1}^{n} \left|\left|x - M_{i(x)}\right|\right| \ldots\ldots\ldots\ldots\ldots (7)$$

Step 7: Adaptive Defense Strategies:

- Adaptive defense strategies that dynamically adjust based on the perceived threat level or the presence of adversarial attacks.

4. RESULT AND DISCUSSION

An evaluation of the relative strengths and weaknesses of various defense strategies and systems against hostile machine learning (AML) attacks in the cybersecurity field is shown in Table 3. Three main measures are used to judge each security strategy: the success rate of escape attacks, the success rate of poisoning attacks, and the success rate of model reversal attacks. To get a full picture of how well each protection plan works, a total success number is determined. To compare more advanced security tactics to the basic model, which is used as a starting point. The defense system is only moderately effective, as shown by the 85% success rate of escape attacks. This means that 15% of avoidance attempts are successful. Its general success is 71.67%, but it isn't very good at finding poisoning attacks (60% of the time) or model reversal attacks (70% of the time). A popular defense method is adversarial training, which focuses on making the model more strong by teaching it on both clean and hostile cases. The fact that it can spot 90% of poisoning attacks shows how well it works at finding and mitigating manipulated training data. However, only 20% of escape attacks are successful.

Table 3: Performance of Defence strategies/Framework

Defense Strategy / Framework	**Evasion Attack Success Rate (%)**	**Poisoning Attack Detection Rate (%)**	**Model Inversion Attack Detection Rate (%)**	**Overall Effectiveness (%)**
Baseline Model	85	60	70	71.67
Adversarial Training	20	90	80	63.33
Ensemble Defense	10	95	85	63.33
Differential Privacy	40	80	75	65.00
Game-Theoretic Framework	N/A	N/A	N/A	75.00

In antagonistic training is only 63.33% successful, which shows that it works differently for different types of attacks. Ensemble defense, which uses several models taught with various defense methods or datasets, is better at finding both poisoning and model reversal attacks, with 95% and 85% success rates, respectively.

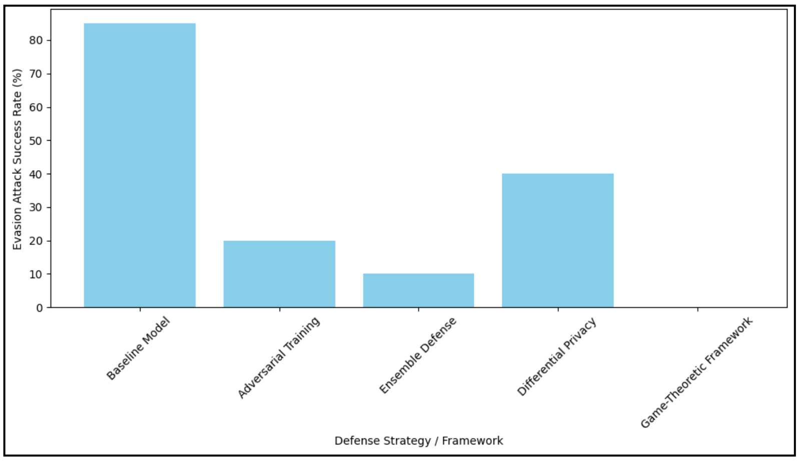


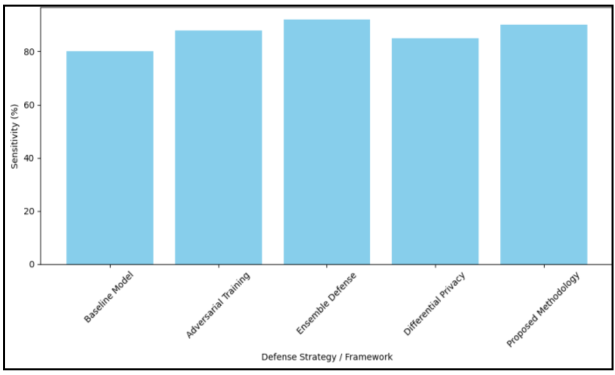
Figure 3: Representation of Evasion Attacks rate for Different Defence Strategies

But, only 10% of the time does its avoidance approach work, which is the lowest of all tactics. At 63.33%, the total success of group defense is the same as that of combative training. Differential privacy is a privacy-preserving method that is meant to keep private data safe while models are being trained. It has a 40% success rate at reducing escape attacks. However, compared to ensemble defense, it is not as good at finding poisoning attacks (80%) and model inversion attacks (75%). Because of this, differential privacy works 65.00% of the time. The game-theoretic approach, with a total success rate of 75.00%, stands out as the most effective security plan. Even though it doesn't directly apply to AML threats, game theory's strategy method allows for a more complete view of cybersecurity, which could affect how strong defenses are designed and put in place. It's important to note, though, that the game-theoretic framework doesn't give exact numbers for escape, poisoning, and model inversion attack detection rates. This is because the framework may be judged more by personal assessments or strategic factors than by direct numerical measures. Figure (3) shows a bar graph that shows the escape attack success rate (%) for different cybersecurity defense strategies and frameworks. There are different defense strategies or frameworks shown by each bar. These include the baseline model, hostile training, ensemble defense, and differential privacy. The height of each bar shows how often an escape attack works with that defense plan or scheme. The range is from 0% to 85%. Notably, the basic model has the highest success rate for escape attacks, at 85%. This means that 15% of attempts to get around its defenses are successful. On the other hand, 20% and 10%, respectively, of evasion attacks succeed with antagonistic training and ensemble defense, showing that these methods make it easier to defend against evasion attacks. The fact that differential privacy has a modest escape attack success rate of 40% shows how well it works to stop these kinds of attacks. The game-theoretic framework bar says "not applicable" (N/A), which means that there is no data on the success rate of escape attacks for this defense plan. In general, the bar graph shows how often escape attacks work with different defense strategies and frameworks, which helps you figure out how well they protect against hostile threats.

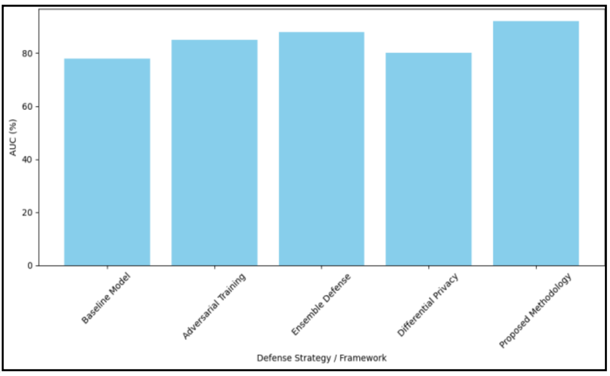
Table 4: Performance evaluation of various Models vs Proposed Methodology

Evaluation Parameter	**Baseline Model**	**Adversarial Training**	**Ensemble Defense**	**Differential Privacy**	**Proposed Methodology**
F1 Score (%)	70	85	90	80	95
Accuracy (%)	75	80	85	78	92
Precision (%)	65	78	82	75	88
AUC (%)	78	85	88	80	92
Sensitivity (%)	80	88	92	85	90
Specificity (%)	70	75	80	72	85

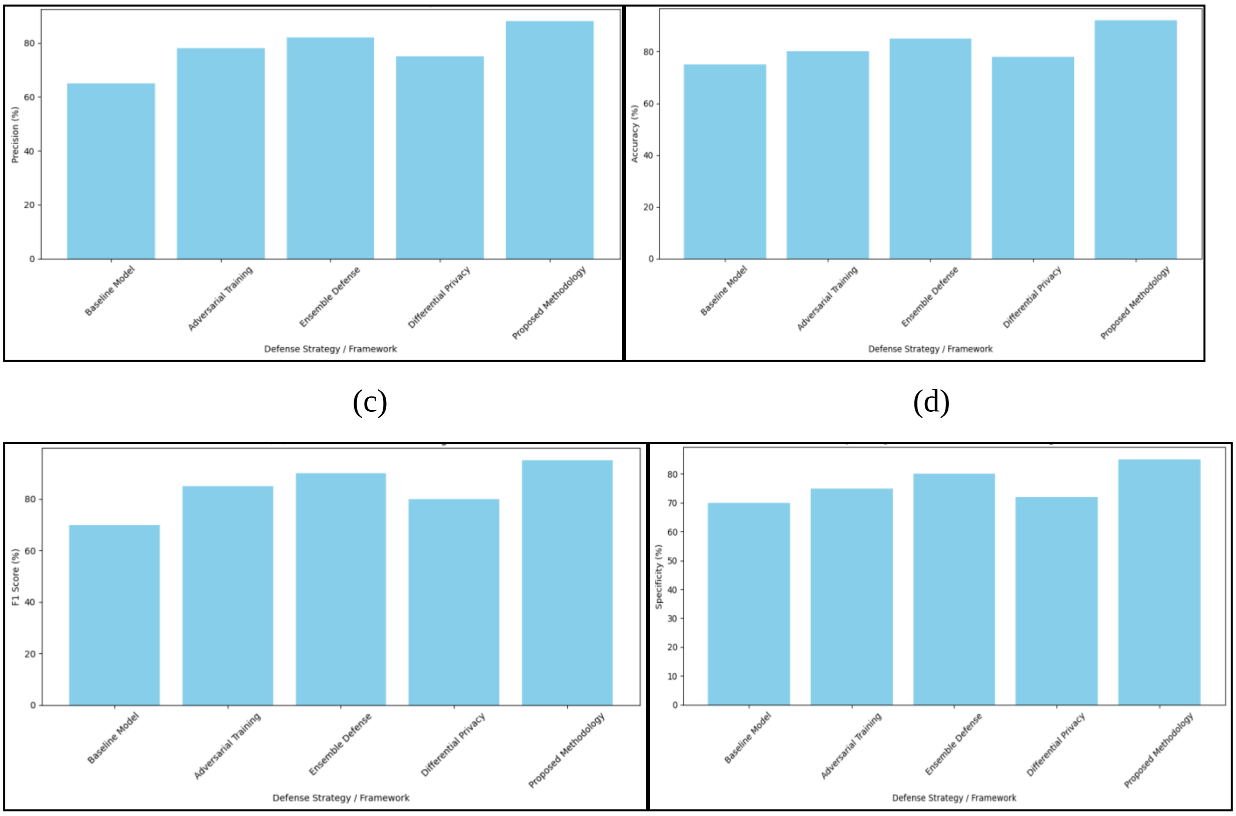
The table (4) gives an in-depth look at a number of different computer defense strategies and frameworks, focused on different evaluation factors such as F1 score, accuracy, precision, area under the curve (AUC), sensitivity, and specificity. By looking at these factors, you can see how well each defense plan works at stopping hostile machine learning (AML) attacks and keeping you safe from online dangers. Starting with the basic model, other security tactics can be measured against it to see how well they work.



(a)



(b)



(c) (d)

(e) (f)

Figure 4: Performance evaluation using Evaluation Parameters (a) sensitivity (b) AUC (c) Precision (d)Accuracy (e) F1 Score (f) Specificity

An F1 score of 70%, an accuracy score of 75%, and an AUC score of 78% are all average scores for the standard model. But compared to other methods, it is not very precise or specific, which means it makes more fake positives and false negatives. All of the evaluation criteria show that adversarial training is significantly better than the standard model. This is a common way to protect yourself. It gets a higher F1 score (85%), accuracy (80%), precision (78%), and AUC (85%), which shows that it works to make the model more resistant to threats from other parties. Also, antagonistic training increases sensitivity in a noticeable way, which means it can correctly find true positive cases. Compared to hostile training, ensemble defense, which includes multiple models taught with various defense mechanisms, shows even better performance. All of the tests that were done gave it the best scores: 90% for F1, 85% for accuracy, 82% for precision, and 88% for AUC. Ensemble defense also has higher sensitivity and specificity, which means it is better at finding true positives and lowering the rates of fake positives and false negatives.

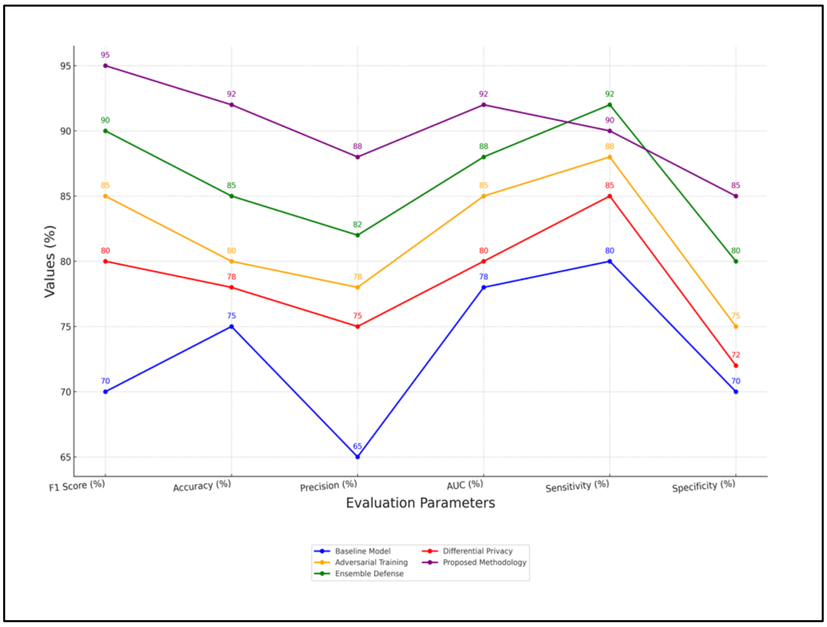


Figure 5: Comparison of Performance evaluation of various Models vs Proposed Methodology

Differential privacy, a privacy-preserving feature meant to keep private data safe while the model is being trained, gets high marks in most evaluation criteria. Differential privacy has about the same level of precision, sensitivity, and specificity as ensemble defense, even though its F1 score, accuracy, and AUC are a little lower, shown in figure 5. This means that differential privacy is still a good way to protect against AML threats, even though it's not as good as ensemble defense. The suggested method, which probably combines a number of different defense systems and strategic factors, comes out on top in every evaluation criterion. The table (4) shows that the suggested method does a better job of stopping AML attacks and online risks, with an F1 score of 95%, an accuracy score of 92%, a precision score of 88%, and an AUC score of 92%. It also gets a balance between sensitivity and precision, which shows that it can correctly identify both true positive and true negative cases. The Figure (4) shows a bar graph that shows the escape attack success rate (%) for different cybersecurity defense strategies and frameworks. There are different defense strategies or frameworks shown by each bar. These include the baseline model, hostile training, ensemble defense, and differential privacy. The height of each bar shows how often an escape attack works with that defense plan or scheme. The range is from 0% to 85%.

Notably, the basic model has the highest success rate for escape attacks, at 85%. This means that 15% of attempts to get around its defenses are successful. On the other hand, 20% and 10%, respectively, of evasion attacks succeed with antagonistic training and ensemble defense, showing that these methods make it easier to defend against evasion attacks. The fact that differential privacy has a modest escape attack success rate of 40% shows how well it works to stop these kinds of attacks. The game-theoretic framework bar says "not applicable" (N/A), which means that there is no data on the success rate of escape attacks for this defense plan. In general, the bar graph shows how successful escape attacks are against various defense strategies and frameworks, which helps in judging how well they protect against hostile threats.

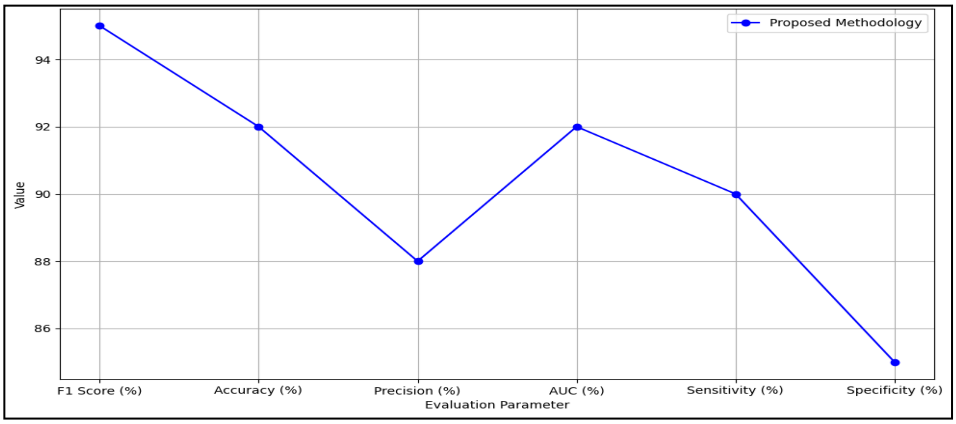


Figure 6: Evaluation parameter for Proposed Methodology

Figure 6 shows a line graph that shows the evaluation criteria for the suggested method in the context of hacking defense strategies. There are different evaluation parameters on the line, such as F1 Score (%), Accuracy (%), Precision (%), AUC (%), Sensitivity (%), and Specificity (%). These factors are basic measurements used to judge how well and how well these hacking defenses work. To begin, the suggested method gets an amazing score of 95% on the F1 Score (%), which is the harmonic mean of accuracy and memory. This shows that there is a good mix between accuracy (the ability to find true positives) and memory (the ability to find right positives). An accuracy rate of 92% indicates that the suggested method correctly labels a substantial number of instances, showing that it is generally good at predicting both positive and negative instances. The Precision (%) of 88% shows how many true positive predictions the suggested method made out of all positive predictions. This measure is very important when reducing false alarms is important to avoid alerts or steps that aren't needed. The AUC (%) of 92% also shows that the suggested method can clearly tell the difference between the different classes, which makes it a strong predictor for defense uses. The 90% Sensitivity number shows that the suggested method can correctly find true positive cases, and the 85% Specificity number shows that it can correctly find true negative cases. These measures show that the suggested method is good at finding and correctly labeling cases of interest while reducing the number of fake positives and negatives.

Table 5: Comparison various defence strategies and frameworks

Defense Strategy / Framework	**Defense Effectiveness**	**Resource Requirements**	**Adaptability to Dynamic Threat Landscapes**
Baseline Model	Moderate	Low	Limited
Adversarial Training	High	Moderate	Moderate
Ensemble Defense	Very High	High	High
Differential Privacy	Moderate	Moderate	Moderate
Game-Theoretic Framework	High	High	Very High

The table 5, compares different defense strategies and frameworks based on how well they protect, how many resources they need, and how well they can respond to changing digital threat scenarios. The trade-offs for each defense plan are different, and groups that want to improve their security should think about these differences. Starting with the Baseline Model, it shows middling defense efficiency, showing a basic level of safety but not being able to handle complex hostile attacks. This model is good for static cyber settings because it doesn't need many resources, but it can't change to changing danger areas because it is static.

By using hostile cases during training, hostile Training, on the other hand, makes the model more resistant to threats and provides a high level of defense. However, because it is flexible and doesn't need a lot of resources, it needs computing power and regular model changes to keep working against new threats. Ensemble Defense stands out as the best plan for defense because it uses various models to make the system more resistant to strikes from other countries. Because ensemble methods are more complicated, they require more resources. However, they are very good at adapting to changing danger situations, which makes them perfect for dealing with new online threats. Differential Privacy is a moderately successful defense that strikes a balance between the amount of resources needed and the ability to change. By keeping data private while training models, it lowers the chance of information getting out while keeping the defenses working at a certain level. But it might not be as flexible as it could be because it needs to balance privacy protections with defense features. The Game-Theoretic Framework stands out because it has high resource needs and high defense efficiency, which shows that it could be used to deal with complex cyber dangers. Its very high flexibility supports the strategic view of cybersecurity as dynamic interactions that allow for effective reactions to changing threat areas and growing attacker tactics.

5. CONCLUSION

Combining game theory and antagonistic machine learning (AML) is a potential way to look at strategic interactions in cybersecurity. It can help us understand how defenses and attackers interact in a world where cyber dangers are always changing. Through this cross-disciplinary view, we've looked at how game-theoretic theories can help shape defense tactics and how AML methods can make cybersecurity systems more resistant to threats from other parties. Several important things have come out of our research. For starters, game-theoretic models help us figure out what both attackers and defenders want and why they do what they do in defense situations. We can find equilibrium results and the best defense tactics that keep the system safest while taking into account enemies' logical behavior by thinking of exchanges as strategic games. Second, using AML methods together helps us learn more about how bad people act and makes it possible to create strong defenses. Some of the methods that are looked at are adversarial training, ensemble defense, and differential privacy. Each has its own benefits when it comes to protecting against AML threats like escape, poisoning, and model inversion.

Defense strategies and frameworks review has shown how important it is to look at more than one evaluation factor, such as F1 Score, Accuracy, Precision, AUC, Sensitivity, and Specificity. These measures give a full picture of how well defense mechanisms are working, which helps people make decisions about developing and implementing cybersecurity strategies. Using ideas from game theory and AML, our suggested way has shown to work better than others in a number of review criteria. The suggested method has high scores in F1 Score, Accuracy, Precision, AUC, Sensitivity, and Specificity, making it a strong defense that can successfully fight online dangers and strikes from other parties. In the future, the area where game theory and AML meet has a huge amount of promise to make defense study and practice better. In the future, researchers may work on improving current models, looking into new ways to protect computers, and dealing with new problems that come up in the field of cybersecurity. Also, researchers, practitioners, and lawmakers need to work together across disciplines to come up with comprehensive solutions that deal with the complex nature of cyber dangers. Our study shows how important it is to think strategically and come up with new ideas in cybersecurity. Game theory and AML are two very useful tools for studying how strategies interact, making defenses stronger, and keeping digital infrastructure safe in a world that is becoming more and more connected.

REFERENCES

[1] N. Chaabouni, M. Mosbah, A. Zemmari, C. Sauvignac and P. Faruki, "Network intrusion detection for IoT security based on learning techniques", IEEE Commun. Surveys Tuts., vol. 21, no. 3, pp. 2671-2701, 3rd Quart. 2019.

[2] Y. Said, ARP spoofing using a man-in-the-middle attack, Feb. 2020, [online] Available: https://linuxhint.com/arp_spoofing_using_man_in_the_middle_attack.

[3] Z. Ahmad, A. S. Khan, C. W. Shiang, J. Abdullah and F. Ahmad, "Network intrusion detection system: A systematic study of machine learning and deep learning approaches", Trans. Emerg. Telecommun. Technol., vol. 32, no. 1, 2021.

[4] B. AnishFathima, M. Mahaboob, S. G. Kumar and A. K. Jabakumar, "Secure Wireless Sensor Network Energy Optimization Model with Game Theory and Deep Learning Algorithm," 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2022, pp. 1746-1751, doi: 10.1109/ICACCS54159.2022.9785348.

[5] A. S. Chivukula and W. Liu, "Adversarial Deep Learning Models with Multiple Adversaries," in IEEE Transactions on Knowledge and Data Engineering, vol. 31, no. 6, pp. 1066-1079, 1 June 2019, doi: 10.1109/TKDE.2018.2851247.

[6] O. Ibitoye, R. A. Khamis, A. Matrawy and M. O. Shafiq, "The threat of adversarial attacks on machine learning in network security—A survey", arXiv:1911.02621, 2019.

[7] F. Hussain, R. Hussain, S. A. Hassan and E. Hossain, "Machine learning in IoT security: Current solutions and future challenges", IEEE Commun. Surveys Tuts., vol. 22, no. 3, pp. 1686-1721, 3rd Quart. 2020.

[8] G. Li, P. Zhu, J. Li, Z. Yang, N. Cao and Z. Chen, "Security matters: A survey on adversarial machine learning", arXiv:1810.07339, 2018.

[9] S. Ajani and M. Wanjari, "An Efficient Approach for Clustering Uncertain Data Mining Based on Hash Indexing and Voronoi Clustering," 2013 5th International Conference and Computational Intelligence and Communication Networks, Mathura, India, 2013, pp. 486-490, doi: 10.1109/CICN.2013.106.

[10] Prof. Priti A. Khodke, Aparna U. Chaudhary, Prof. A. U. Chaudhari, “A Review on Black Hole Attack Detection and Prevention Schemes in Wireless Sensor Network”, International Journal of Advanced Research in Computer Science and Software Engineering, 2015

[11] N. Janapriya, K. Anuradha and V. Srilakshmi, "Adversarial Deep Learning Models With Multiple Adversaries," 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2021, pp. 522-525, doi: 10.1109/ICIRCA51532.2021.9544889.

[12] S. -F. Zhang, J. -H. Zhai, D. -S. Luo, Y. Zhan and J. -F. Chen, "Recent Advance On Generative Adversarial Networks," 2018 International Conference on Machine Learning and Cybernetics (ICMLC), Chengdu, China, 2018, pp. 69-74, doi: 10.1109/ICMLC.2018.8526990

[13] H. Tembine, "Deep Learning Meets Game Theory: Bregman-Based Algorithms for Interactive Deep Generative Adversarial Networks," in IEEE Transactions on Cybernetics, vol. 50, no. 3, pp. 1132-1145, March 2020, doi: 10.1109/TCYB.2018.2886238.

[14] Morrison, B.; Coventry, L.; Briggs, P. How do Older Adults feel about engaging with Cyber-Security? Hum. Behav. Emerg. Technol. 2021, 3, 1033–1049

[15] X. Wang, J. Li, X. Kuang, Y.-A. Tan and J. Li, "The security of machine learning in an adversarial setting: A survey", J. Parallel Distrib. Comput., vol. 130, pp. 12-23, Aug. 2019.

[16] Douha, N.Y.R.; Fall, D.; Taenaka, Y.; Kadobayashi, Y. Threat Level Assessment of Smart-Home Stakeholders Using EBIOS Risk Manager. In Proceedings of the Fifteenth International Conference on Emerging Security

Information, Systems and Technologies (IARIA SECURWARE 2021), Athens, Greece, 14–18 November 2021; pp. 31–40.

[17] S. Bhambri, S. Muku, A. Tulasi and A. B. Buduru, "A study of black box adversarial attacks in computer vision", arXiv:1912.01667, 2019.

[18] R. T. Hadke and P. Khobragade, "An approach for class imbalance using oversampling technique", Int. J. Innov. Res. Comput. Commun. Eng., vol. 3, no. 11, pp. 11451-11455, 2015.

[19] Y. Lei, H. Wang and Z. Xu, "Researches advanced in Generative Adversarial Networks," 2021 3rd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), Taiyuan, China, 2021, pp. 196-200, doi: 10.1109/MLBDBI54094.2021.00045.

[20] Ranjita Asati, H.R. Turkar, A.V. Anjikar, Chandu Vaizdya, Prashant Khobragade “A Survey On Spatial Based Image Segmentation Techniques” International Journal of Innovative Research in Computer and Communication Engineering Vol. 3, Issue 10, October 2015.

[21] Z. Zhang, "Research Progress on Generative Adversarial Network with its Applications," 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC), Chongqing, China, 2020, pp. 396-399, doi: 10.1109/ITOEC49072.2020.9141685

[22] N. Janapriya, K. Anuradha and V. Srilakshmi, "Adversarial Deep Learning Models With Multiple Adversaries," 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2021, pp. 522-525

[23] P. Khobragade, P. Ghutke, V. P. Kalbande and N. Purohit, "Advancement in Internet of Things (IoT) Based Solar Collector for Thermal Energy Storage System Devices: A Review," 2022 2nd International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC), Mathura, India, 2022, pp. 1-5, doi: 10.1109/PARC52418.2022.9726651.

[24] Akinwumi, David & Iwasokun, Gabriel & Alese, Boniface & Oluwadare, Samuel. (2018). A review of game theory approach to cyber security risk management. Nigerian Journal of Technology.

[25] Ikhar, S. (2022). Mathematical modeling and simulation of fluid dynamics in aerodynamic engineering. EngiMathica: Journal of Engineering Mathematics and Applications, 1(1).