

---

## ANALYSES OF ARTIFICIAL INTELLIGENCE APPLICATIONS IN CYBERSECURITY

Ani A.A.<sup>1</sup>, Dimson I.C.<sup>2</sup>, Obi M.C.<sup>3</sup>

1: Dept. of Computer Engineering, Madonna University of Nigeria, Akpugo Campus, Enugu State, Nigeria.

2: Dept. of Electronic and Computer Engineering, Nnamdi Azikiwe University, Awka, Nigeria.

3: Dept. of Computer Engineering, Caritas University, Enugu, Nigeria.

### ABSTRACT

In the face of growing cyber-attacks, the effectiveness of artificial intelligence (AI) in the field of cybersecurity has become a crucial breakthrough. The revolutionary influence of AI in boosting cybersecurity strategies' efficacy is examined in this paper. This paper emphasizes the efficiency that AI provides to the cybersecurity environment by examining the ways it optimizes threat detection, prevention, and incident response. The evaluation also noted the difficulties and concerns related to AI integration, such as possible prejudices and hostile actions (i.e., adversarial attacks). Organizations stand to transform cybersecurity methods, strengthening their defenses and influencing a safer digital future as they increasingly utilize AI's capabilities. This work analyzes the performances and provides recommendations for various artificial intelligence models as it relates to cybersecurity.

**KEYWORDS/PHRASES:** Cybersecurity, artificial intelligence, machine learning, cyber-attacks

### 1. INTRODUCTION

The rate of ongoing technological advancements has led to greater interconnectivity in today's world and as time progresses the usage of these digital devices is steadily increasing. Research studies in fields of computer science and engineering with regards to artificial intelligence (AI), cloud computing and cyber security is greatly facilitated with the growth in the cyberspace. Different industries, especially health and financial institutions, have employed the use of these technologies in a variety of ways such as in the use of wearable devices in monitoring healthcare (Nahavandi et al., 2022).

This has led to an onslaught of data being generated every second as a result of the increase in the network traffic.

The availability of all these data makes these organizations and their users prone to cyber threats and attacks. Cyber-attacks often lead to economic damages which can be devastating for its victims. It can often lead to exposure of personal information to dangerous entities (Lallie et al., 2021). According to Sharif and Mohammed (2022), cybercrime was estimated to have risen up by 600% since the COVID-19 pandemic and ransomware would cost \$10.5 trillion annually by 2025. Based on the data from Sharif and Mohammed (2022) and Zeadally et al. (2020), some of the prevalent cyber threats include:

1. **Phishing attacks:** An act of sending communications that mimics a reliable source, can be often through email or SMS messages. It is usually aimed at obtaining tricking the victim to give out sensitive information like bank account details or login details to accounts.
2. **Denial of Service (DoS) attacks:** A form of attack aimed at overburdening a system's resources by flooding it with traffic. This disables the system making it unable to work normally. A variant of these attack by using multiple attack devices is termed Distributed Denial of Service (DDoS).
3. **Malware:** This is concerned with the installation of unwanted and malicious software on the victim's mobile device or system. It is often embedded in unknown and insecure web links or file attachments.

4. **Man – in – the – middle (MiTM) attacks:** This occurs when a perpetrator intercepts messages between individuals and continues the communication with the hope of extracting information. This type of attack often occurs virtually by session hijacking. This happens when the victim tries to access confidential documents or perform a classified process on a public network, thus enabling the attacker to hijack the IP address and hijack the session.

The types of cyber threats are not limited to the above and it is rapidly evolving as the attackers are changing and modifying their methods as time passes. This has prompted researchers to develop means to subvert these threats giving rise to cybersecurity. Cybersecurity is concerned with the application of various techniques and processes in the quest to maintain data integrity and protect systems from malicious attacks (AL-Hawamleh, 2023; Shaukat, Luo, Varadharajan, Hameed, & Xu, 2020). Conventional methods of cyber threat detection involved the use of antivirus software, firewalls, use of encryption software amongst others (Zeadally et al., 2020). These methods offered some accurate detections but they were not dynamic enough for the ever-changing nature of cyber-attacks. They were lacking due to strict design and the required user involvement. Thus, spurring the need of employing automatic means of monitoring cyber threats, producing more efficient detections for the new types of threats and predicting future trends.

In recent times, research on AI-based cybersecurity techniques has solved this problem by providing dynamic models capable of adapting to different threats detected in a network so long as there is appropriate training. AI has become a thriving research area in cybersecurity due to its ability to adapt to different scenarios and the availability of huge datasets for training the models. Machine Learning (ML), Deep Learning (DL) and Natural Language Processing (NLP) are subfields of AI which is commonly employed in the design of various cybersecurity technologies

as can be seen in the products of security organizations like the Cognito (Vectra, 2023), Intercept X (Sophos, 2023) and Enterprise Cloud (Broadcom, 2023). Figure 1 below shows the relationship between AI, ML, DL and NLP fields.

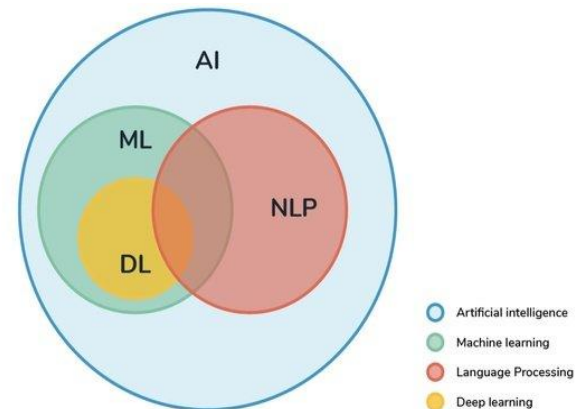


Figure 1: The interrelationship between various AI focus areas

There has been a lot of research and review papers focusing on the application of various AI technologies in the detection of cyber-attacks. However, this paper will cover the relevance of AI research in cybersecurity by reviewing related literature on AI related cybersecurity techniques with focus on Machine Learning, Deep Learning and Natural Language Processing models. It also looked into performance analyses of different models, as it relates to detection, prevention and prediction of various cyber threats.

## II. LITERATURE REVIEW

### 2.1 Use of AI in Cybersecurity

Conventional cyber threat detection involves security control by the user in terms of scheduled software and antivirus updates, anomaly and signature-based detections (Snehi et al., 2021) and game theory (Pawlick et al., 2019). These methods had some limited capabilities because they were dependent on the user or on the current pattern of the threat. In order to eliminate and reduce these difficulties, the application of AI-based methods thrived.

Artificial Intelligence (AI) is a field of study that involves the process of enabling machines or systems to think intelligently and make

deductions based on previous information. Due to the sophistication of cyber-attacks and the need in ensuring a secure network, ML, DL and NLP, the sub-fields of AI posed great interest to researchers and educators. Although, these three terms are synonymously used in publications to represent any AI-based activity, they differ in the description and implementation processes.

As can be seen from Figure 1, AI is the area which encompasses all the other fields. It deals with any form of automation to the machine to make it think and make deductions intelligently. Meanwhile, ML is a form of AI that utilizes learning algorithms to fulfilling such automated tasks without being explicitly told to do so. Deep learning, which is carved out of ML, deals with expanding the learning scope of ML using the implementation of deep neural networks (i.e., neural networks with two or more hidden layers). NLP is a technique born out of ML and DL algorithms in order to train a system to understand, predict and process human text patterns.

AI techniques have been prominent in the past decade in connection with researches based on network intrusion detection, Internet of Things (IoT) security, threat prevention and prediction and in other areas connected with cybersecurity. Shaukat et al. (2020) presented a review of the performance of various ML techniques over a decade. Their analysis focused on the technique utilized only in relation to malware, spam and intrusion detection. Moreover, in their evaluated each model as it relates to popular datasets.

Ravi et al. (2021) reviewed the various deep learning methods, trends and applications as it relates to some cyber threats such as phishing, malware, spam and botnet detection. They also discussed cybersecurity as it relates to the application of key concepts like blockchain and natural language processing. It asserted that DL is crucial to cybersecurity tasks but that research on these techniques is still at its infancy.

With a focus on mobile network security, Gupta et al. (2022) analyzed the threats found in mobile devices such as unauthorized accesses and fraudulent links and reviewed the relevance of using an AI-based model for accurate detection and security. Their paper discussed related ML and DL techniques applied for various cyber threats detection. Meanwhile, Tojiboyev (2023) presented a review on the impact AI has made on the cybersecurity field.

Other research papers proposed custom models for cyber threat detection as it relates to IoT system (Banaamah & Ahmad, 2022; Ghillani, 2022), smart grids (Berghout et al., 2022) and mobile devices (Rodriguez et al., 2021)

## 2.2 AI Machine Learning

Conventional machine learning algorithms used in related research can be classified into three types: Supervised, Unsupervised and Reinforcement algorithms.

Supervised algorithms classify the threats based on predefined classes in the trained dataset. It includes methods like Support Vector Machine (SVM), Naïve Bayes (NB), Random Forest (RF) and Decision tree (DT). On the other hand, unsupervised algorithms have no predefined classes and it attempts to find similar patterns in the dataset. Unsupervised models commonly applied in cybersecurity include K-means clustering amongst others. The third class of ML models, Reinforcement learning, deals with training the data based on the environment. Different possibilities are analyzed before an action is taken in order to figure out if it solves the problem at hand before implementation. It was a focus in the research by Nguyen and Reddi (2021).

The description of some commonly used cybersecurity AI Algorithms can be seen in Table 1 below. It also details the benefits each model offers in their various implementations.

**Table 1: Description of commonly used ML models**

| ML algorithm              | Description   | Benefits  | Challenges observed  |
|---------------------------|---|---|--|
| DT                        | A classification ML algorithm that models tasks and their possible outcomes in a tree-like structure  | <ul style="list-style-type: none"> <li>• Ease in understanding decisions made by the model</li> <li>• Can handle different data (numerical and categorical) without prior processing</li> </ul> | <ul style="list-style-type: none"> <li>• Prone to overfitting</li> <li>• Poor adaptation to high-dimensional datasets</li> </ul> |
| RF                        | A cluster of decision trees that enhances accuracy and minimizes overfitting  | Accuracy improvement  | Computationally expensive  |
| NB                        | This is a method often utilized in text classification and sentiment analysis. It works on a probability-based principle that assumes that features are independent.      | Efficient for tasks like mail spam detection  | Makes assumptions which does not scale well in real-world scenarios  |
| SVM                       | SVM methods finds the best hyperplane that separates classes of data.   | Ability to handle high-dimensional data   | Requires careful tuning of hyperparameters   |
| K-Nearest neighbors (KNN) | This is a classification method that group instances based on their k-nearest neighbors.  | Able to perform intuitive search for alike patterns in network traffic  | <ul style="list-style-type: none"> <li>• Computationally expensive for large datasets</li> <li>• Unscalable</li> </ul>           |
| CNN                       | This is a DL model well-suited for analyzing visual data  | Used in capturing visual anomalies in the data  | Requires extensive tuning for optimal performance  |
| DBN                       | This is a DL model based on layered Restricted Boltzmann machines. It is designed to learn sequential data representations.   | Able to detect anomalies in the network traffic   | Unable to transfer training to different dataset   |
| AE                        | This is an unsupervised learning technique that encodes the input data into a suitable representation. After efficient learning, it decodes it back to its original state | <ul style="list-style-type: none"> <li>• Able to learn high-dimensional data</li> <li>• Useful in feature learning of the dataset</li> </ul>  | Sensitive to data quality and requires preprocessing.  |

Mihoub et al. (2022) proposed a ML model for the detection of DoS/DDoS attacks based on the Random Forest algorithm. Meanwhile, Trivedi et al. (2020) presented a comparative analysis of various ML algorithms as it relates to credit card fraud detection.

ML-based implementations may often include a combination of different methods in order to enhance performance, this is termed *ensemble learning*. The research by (Sarnovsky & Paralic, 2020) presented an ensemble approach by utilized NB, DT and RF in the detection of DoS

attacks on the NSL-KDD dataset. Another typical instance of the ensemble learning approach is seen in the paper by Zuhair et al. (2020) for zero-day malware detection.

It should also be noted that when a previously trained model is applied for training another dataset for another task, this is known as *transfer learning* as can be seen in the research by (Sarker, 2021)

It is worth noting that quite a number of research has been focused on the implementation of cybersecurity techniques based on ML methods in the past few years. This is because the basis of cyber threat detection is on classification. A ML model is able to classify unknown threats into the right category based its degree of accuracy. Due to the limit in performance of some ML models, DL models were utilized for robust adaptation and utilization of more parameters in the learning process.

### 2.3 Deep Learning

The emergence of DL models in cybersecurity research can be attributed to the improvement of processing capacity of present-day computer systems because DL algorithms are a computationally intensive learning and would require a lot of data and processing power for sufficient training of the model (Fan et al., 2021). In order to obtain higher predictor models and utilize more hyperparameters for detections, DL-based methods started being applied in the design of various cybersecurity techniques as can be seen in Table 4. From related literature, DL models commonly utilized are the Convolutional Neural Networks (CNN), Autoencoders (AEs), and Deep Belief Networks (DBNs).

The paper by Tran et al. (2022) presented a custom artificial neural network with 4 hidden layers which was developed to validate the cutting signal of a CNC machine. This model was able to distinguish between a real signal and a fake signal that attempts to alter the working of the machine. Meanwhile, Rhode et al. (2018) utilized the RNN algorithm for early detection of malware on systems. This model showed a satisfactory performance of 96% accuracy.

Some researchers presented an ensemble approach of DL methods in the detection of zero-day ransomware attacks (Zahoor et al., 2022) and fake links (phishing) (Aldakheel et al., 2023)

The lack of data needed for appropriate tuning of the hyperparameters of the DL models led researchers to develop their own data. This was achieved by means of generative DL models such as the Generative Adversarial Network (GAN). This is a DL model that learns the patterns in the data and creates it own input from it, afterwards it tries to discriminate between the two inputs. The discrimination process is continuous until the network is unable to distinguish the real data from the fake data. This is a promising area for cybersecurity because GANs could help in generating a sort of cyber attacks that seems realistic and can be used for training other models to maximize performance (Samtani et al., 2020)

### 2.4 Natural Language Processing

NLP can be used in the prediction and immediate detection of cyber threats. It involves the application of ML-based techniques to develop tools that can monitor the network of unstructured data (emails, SMS) for suspicious activity. This also includes data from websites, system logs and online messages. The research in this area can be seen from the paper by (Thapa, 2022) which proposed the implementation of VADER (Valence Aware Dictionary For Sentiment Reasoning) in the classification of Twitter and Reddit cybersecurity dataset. The research also involved human participation in the classification. When the results of the two models were calibrated, the NLP showed a lower precision rate of 86% against the 100% of human rating. This points out the need for further research and for bigger datasets for appropriate training of the NLP model.

### 2.5 Metrics for Assessing AI Models

The performance of ML and DL models are based on certain set of criteria(metrics). These metrics determine the optimal performance of the model. Before delving into the common

metrics, a basic understanding of terms such as True Positive (TP), False Positive (FP), False Negative (FN) and True Negative (TN) is crucial in understanding the formulas relating to the ML performance metrics.

True Positive represents all the instances which the model correctly predicts that is true while

False positive represents those instances incorrectly predicted as true. Likewise, True Negative represents the instances correctly predicted as untrue while False negative are those classes incorrectly labelled untrue.

**Table 2: Confusion matrix**

|                |                 | Actual Data        |                    |
|----------------|-----------------|--------------------|--------------------|
|                |                 | Predicted Positive | Predicted Negative |
| Predicted Data | Actual Positive | TP                 | FN                 |
|                | Actual Negative | FP                 | TN                 |

The diagram above represents a typical confusion matrix generated after the successful training and testing of the ML model. It is used in analyzing parameters such as:

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

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

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

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

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

The accuracy specifies the proportion of correctly predicted occurrences to the total occurrences. Precision indicates the proportion of correctly predicted positive chances out of all predicted positives. Recall (also referred to as the true positive rate) measures the proportion of correctly predicted positive chances out of all actual positives. Specificity (also referred to as the true negative rate) measures the proportion

of correctly predicted negative occurrences out of all actual negatives. F1-Score is the harmonic mean of precision and recall.

### III. PERFORMANCE ANALYSES

Due to an ever-changing nature in the technological world, there is need to ensure a secure communication channel. The advent of AI-based techniques in cybersecurity has provided immediate and beneficial solutions in the detection and elimination of cyber threats. Tables 3 and 5 outlined some the relevant applications of various AI schemes in the technological sphere. The choice of model for each model varied depending on the research aim, for instance, in order to evaluate the efficiency of a DL-model for botnet detection Ahmed et al. (2022) proposed a feed-forward propagation ANN model that first extracts the feature data before training. The ANN model consists of a varying number of hidden layers and it employed the Adam optimization process by TensorFlow so as to achieve efficient computations. After adequate testing, this custom model achieved an accuracy value of 99.2%. The overview of various researches related to botnet detection using Deep Learning models can be found at Sarker (2021).

**Table 3: Summary of the Performance of ML-Based Models Based on Related Literature**

| Reference                   | ML training model   | Research focus  | Dataset used                           | Performance  |
|-----------------------------|---|---|--|--|
| (Alharbi & Alsubhi, 2021)   | ExtraTrees classifier (with Pearsons correlation-features)                      | Detection of botnet traffic                           | CTU-13 and IoT-23                      | Achieved 99% accuracy and 100% precision and recall  |
| (Chohan et al., 2023)       | Linear SVM, Ada Boost, AE, Multilayer Perceptron                                | Intrusion detection system                            | UNSW-NB15                              | Ada Boost achieved is 98.3% accuracy   |
| (Sarnovsky & Paralic, 2020) | NB, DT, RF  | Detection of DoS attacks                              | NSL-KDD                                | Achieved accuracy and precision of 99.80%  |
| (Syed et al., 2020)         | NB-based classifier, DT-based classifier, MLP                                   | DoS attack detection on IoT networks                  | Custom                                 | The NB-based classifier achieved the maximum accuracy of 99.9%   |
| (Mihoub et al., 2022)       | RF  | DoS/DDoS detection and mitigation                     | Custom Bot-IoT dataset                 | Accuracy of 99.81%   |
| (Outman et al., 2023)       | KNN, LSTM, SVM- and DT-based methods, Isolation Forest                          | Detection of MitM attacks on process control networks | Realtime SCADA datasets                | Coarse Tree. A DT-based model achieved the optimal performance with 100% accuracy at 0.45ms training time. |
| (Zuhair et al., 2020)       | DT, NB  | Zero-day attack detection (malware)                   | Custom                                 | Achieved 97% accuracy  |
| (Usman et al., 2021)        | SVM, DT, NB, MBK  | Malware detection in IP                               | Custom                                 | SVM achieved a high precision rate of 98%  |
| (Trivedi et al., 2020)      | Comparative analysis of various ML methods                                      | Credit card fraud detection                           | Custom                                 | RF displayed the maximal accuracy of 95.988%   |
| (Brindha et al., 2023)      | GRU   | Fake mail detection                                   | Enron                                  | Achieved 99.72% accuracy   |
| (Gangavarapu et al., 2020)  | Ensemble learning (PCA, RF, SVM, NB)  | Spam mail detection                                   | Custom                                 | RF achieved the highest accuracy of 93%  |
| (Y. Wei & Sekiya, 2022)     | Comparative analysis of common ML models (K-means clustering, SVM, NB, KNN, RF) | Phishing detection                                    | UCI_2015, MENDELEY_2018, MENDELEY_2020 | RF showed a maximum accuracy of 96.84% for a large dataset and 96.94% for a smaller dataset.               |

For IoT systems, network threats such as DoS and intruder monitoring abound. This has prompted a lot of related research for Intrusion Detection systems using various ML methods (Ahmad et al., 2021; Snehi et al., 2021). NB, DT

and RF pose as some of the commonly applied models and they achieved high accuracy figures summarized in Table 3.

The detection of phishing and spam threats in the cyberspace cannot be overlooked as

researchers such as Guo et al. (2021) proposed an adaptive spam detection model which made use of neural networks. This model titled called Co-Spam utilized a Bi-directional Autoencoder (Bi-AE) for modelling the pattern characteristics, a graph convolutional network (GCN) for learning the encodings in the features and the Long Short-Term memory (LSTM) for recalling all the learnt features. Their research showed a significant precision rating of 94.32% on the Twitter dataset and a 5% difference in rating when compared with other models. A review of related research on spam and phishing detection

can be found in (Gangavarapu et al., 2020; Ravi et al., 2021; Shaukat et al., 2020)

A lot of datasets are used for effectively training and testing of various cyber threats. Often, a study may use a web scraping tool to generate their own dataset or combine two or more datasets for a more robust test base. Table 4 below lists some of the frequent datasets that can be found in related literature. It also details their unique characterizations.

Table 4: Common datasets utilized for training ML models(Gupta et al., 2022; Muzaffar et al., 2022; Shaukat et al., 2020)

| Dataset Name           | Attributes  | Year        |
|------------------------|---|-------------|
| NSL-KDD                | All sorts of cyber threats                                    | 2009        |
| Enron                  | Email spam  | 2015        |
| CTU-13                 | Botnet traffic  | 2011        |
| DARPA                  | IDS based dataset   | 1998        |
| F-Droid                | Online marketplace for open-source apps                       | -           |
| DREBIN                 | Android malware dataset                                       | 2010 - 2012 |
| VirusShare, VirusTotal | Malware apps for different OS                                 | -           |
| CAIDA '07              | DDoS traffic data collected in                                | 2007        |
| Alexa top sites        | Malicious domain names  | -           |
| Bot-IoT                | Simulated traffic on IoT networks (DoS, DDoS, and keylogging) | -           |

For analysis on bigger datasets that is usually a combination of two or more of the common datasets or a custom dataset pulled together by the researchers, the use of DL models shows greater performance as can be seen in Table 5. DL models are characterized by their high-

performance rates as a result of their learning process which is facilitated by the processing of data through multiple layers. Several of these models can be applied for different tasks such as the use of CNN for phishing and malware detections (Sarker, 2021).

Table 5: Summary of the performance of DL-based models based on related literature

| Reference            | DL training model                 | Research focus   | Dataset used | Performance   |
|----------------------|-----------------------------------|--|--------------|---|
| (Tran et al., 2022)  | Custom model with 4 hidden layers | Development of a DL-based model interfaced with an IoT device that validates the cutting signals of a CNC machine. | Custom       | Obtained 100% on the precision, recall and F1-score |
| (Ahmed et al., 2022) | Backpropagation and deep neural   | Development of a model for efficient   | CTU-13       | Achieved 99.25% accuracy                            |



|                          |  |   |  |  |
|--------------------------|--|---|--|--|
|                          | network with varying layers for the training and testing phases          | Botnet attack detection                                 |  |  |
| (Zahoor et al., 2022)    | Ensemble learning (Attribute learning-based Deep Contractive AE and KNN) | Zero-day ransomware attacks                             | Custom                                 | Demonstrated 95% recall and 92.8% accuracy   |
| (Rhode et al., 2018)     | Recurrent Neural Networks (RNN)  | Early malware detection                                 | Custom                                 | Achieved max accuracy of 96%   |
| (Azmoodeh et al., 2019)  | Hybrid methodology incorporating CNN                                     | IoT-based malware detection of sensitive infrastructure | Custom                                 | Achieved an accuracy of 99.68%, precision and recall values of 98% approx.                   |
| (Y. Wei & Sekiya, 2022)  | Comparative analysis of popular DL models like CNN, FCNN and LSTM        | Phishing detection                                      | UCI_2015, MENDELEY_2018, MENDELEY_2020 | RF showed a maximum accuracy of 87.99% for a large dataset and 91.38% for a smaller dataset. |
| (B. Wei et al., 2019)    | CNN  | Fake URL detection (phishing)                           | Custom                                 | Achieved a true detection rate of 86.63%   |
| (Aldakheel et al., 2023) | Ensemble learning (CNN + RF)   | Fake URL detection                                      | PhishTank dataset + custom             | Achieved accuracy of 98.77% and precision of 8.01%   |

#### IV. CHALLENGES AND FUTURE PROSPECTS

The field of AI research is greatly progressing as can be seen in the sections above, but there still exists several limitations to its application for effective detection of cyber threats and attacks. In this section, we will discuss the various challenges by researches and the suggested areas of future research.

1. **Dated datasets:** As can be seen from the list of commonly used datasets used for training the AI models, they are dated as far back as 1998. Thus, they are outdated and requires some new data based on the current threats in the cyberspace. This is a great issue because AI models require a huge volume of

data for efficient training and this is lacking in the currently available data.

2. **Lack of clear interpretation of performance:** There is the absence of a specified set of metrics in defining the performance of an AI methodology. It brings about the question of the best way of categorizing and comparing various methods. Thus, it would be useful to have an industry defined standard for measuring how well a model works for particular tasks.
3. **Prone to adversarial attacks:** The recent trend in deep learning approaches gives rise to adversarial processes. Adversarial processes usually involve the use of deep neural networks in

creating mimics that can deceive a properly trained AI model in making incorrect classification. A review of adversarial attacks and defenses on AI-based systems can be found at (Chakraborty et al., 2021; Puttagunta et al., 2023)

4. **Privacy issues:** In the quest to obtain the training data care is to be taken to protect user information. This serves as a huge setback for researches as data is not readily available so as to reduce the risk of adversarial attacks. Therefore, it is essential to develop a privacy-enabling technology that does not encroach on performance.

## V. CONCLUSION

As cybersecurity presents a great concern in recent cloud-based technologies such as IoT system, wearable devices and in the use of social media platforms, it is paramount to have effective techniques to deter any malicious threats and attacks. There exists some conventional means of detection but the advancements in artificial intelligence provides

an avenue to automate these detection means and possibly provide real-time detection.

This study reviewed the efficiency and relevance of the use of AI-based methods in processes of cybercrime identification and monitoring. It provides an analysis of the performance of the various AI models based on different cyber threats and datasets for testing. This showed NB, DT and RF as the recommended machine learning algorithms since they show high accuracy rates. Meanwhile, while implementing a deep learning algorithm, especially on bigger datasets, a deep neural network model such as CNN is recommended as this is able to learn the patterns in the data and achieve a high accuracy rate. Afterwards, the challenges as it relates to the implementation of AI-based models in cybersecurity were discussed and the future prospects in these areas were outlined.

This paper should serve as a guide for future researchers and students who seek to understand the relevance of artificial intelligence in the field of cybersecurity and how it can be implemented in order to achieve an accurate and precise detection model.

## REFERENCES

- Ahmed, A. A., Jabbar, W. A., Sadiq, A. S., & Patel, H. (2022). Deep learning-based classification model for botnet attack detection. *Journal of Ambient Intelligence and Humanized Computing*, 13(7), 3457–3466. <https://doi.org/10.1007/s12652-020-01848-9>
- AL-Hawamleh, A. M. (2023). Predictions of Cybersecurity Experts on Future Cyber-Attacks and Related Cybersecurity Measures. *International Journal of Advanced Computer Science and Applications*, 14(2), 801–809. <https://doi.org/10.14569/IJACSA.2023.0140292>
- Banaamah, A. M., & Ahmad, I. (2022). Intrusion Detection in IoT Using Deep Learning. *Sensors*, 22(21). <https://doi.org/10.3390/s22218417>
- Broadcom. (2023). *Symantec Enterprise Cloud*. Broadcom. <https://www.broadcom.com/products/cyber-security>
- Fan, J., Ma, C., & Zhong, Y. (2021). *A selective overview of deep learning*. 36(2), 264–290. <https://doi.org/10.1214/20-sts783.A>
- Ghillani, D. (2022). Deep Learning and Artificial Intelligence Framework to Improve the Cyber Security. *American Journal of Artificial Intelligence*. <https://doi.org/10.22541/au.166379475.54266021/v1>
- Nguyen, T. T., & Reddi, V. J. (2021). Deep Reinforcement Learning for Cyber Security. *IEEE Transactions on Neural Networks and Learning Systems*, 32(1), 870. <https://doi.org/10.1109/TNNLS.2021.3121870>
- Pawlick, J., Colbert, E., & Zhu, Q. (2019). A game-theoretic taxonomy and survey of defensive deception for cybersecurity and privacy. *ACM Computing Surveys*, 52(4), 1–28. <https://doi.org/10.1145/3337772>

- 
- Samtani, S., Kantarcioglu, M., & Chen, H. (2020). Trailblazing the Artificial Intelligence for Cybersecurity Discipline: A Multi-Disciplinary Research Roadmap. *ACM Transactions on Management Information Systems*, 11(4).  
<https://doi.org/10.1145/3430360>
- Sarker, I. H. (2021). Deep Cybersecurity: A Comprehensive Overview from Neural Network and Deep Learning Perspective. *SN Computer Science*, 2(3).  
<https://doi.org/10.1007/s42979-021-00535-6>
- Sophos. (2023). *Sophos Endpoint Protection: Intercept X with EDR*.  
<https://www.sophos.com/en-us/products/endpoint-antivirus.aspx>
- Syed, N. F., Baig, Z., Ibrahim, A., & Valli, C. (2020). Denial of service attack detection through machine learning for the IoT. *Journal of Information and Telecommunication*, 4(4), 482–503.  
<https://doi.org/10.1080/24751839.2020.1767484>
- Thapa, B. (2022). *Sentiment Analysis of Cyber Security Content on Twitter and Reddit*. 95–107.  
<https://doi.org/10.5121/csit.2022.120708>
- Tojiboyev, I. (2023). *The Influence and Limitations of AI in Cybersecurity Domain*. 18, 53–59.