JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN CYBERSECURITY THREAT DETECTION

-Dr. Devendra Singh, Professor, Haryana Institute of Public Administration

ABSTRACT

The method in which we identify, prevent, and respond to emerging cyber threats has been revolutionised by the use of artificial intelligence (AI), which has become an essential component of current cyber security systems. In response to the increasing complexity and number of cyber attacks, traditional security approaches frequently fail to keep up. This has led to the development of AI-driven solutions, which provide better capabilities in threat identification and mitigation. Machine learning (ML) and deep learning algorithms are utilised by AI-powered systems in order to analyse large volumes of data, recognise patterns, and find anomalies that may suggest possible security breaches. These systems are able to monitor and analyse network traffic, user behaviours, and other data sources in a very short amount of time. This allows them to identify complex threats such as zero-day assaults, advanced persistent threats (APTs), and ransomware, which would otherwise be undiscovered by traditional methods. Artificial intelligence's capacity to continually learn and adapt to new attack vectors is one of the most significant benefits it offers in the field of cyber security. The detection accuracy of artificial intelligence systems may be improved over time by the use of past data and real-time monitoring. This allows these systems to become increasingly successful in recognising new threats. Not only does this proactive strategy increase the rate at which risks are identified, but it also lessens the need on human intervention, which may be both time-consuming and prone to errors. In addition, artificial intelligence has the capability to automate monotonous security chores, such as vulnerability detection and patch administration, which frees up cyber security specialists to concentrate on more complicated problems. Artificial intelligence has the potential to assist with threat intelligence by correlating data from a variety of sources and generating insights that may be put into action to inform decision-making. Organisations are able to construct cyber security frameworks that are more robust and resilient, and that are able to respond to attacks in real time, by integrating artificial intelligence with their existing security infrastructure. As cyber threats continue to change, it is anticipated that artificial intelligence will play an increasingly important role in cybersecurity. This is because continual improvements in AI algorithms and technology will enable defence mechanisms that are more complex and effective. This integration of artificial intelligence in cybersecurity marks a paradigm leap, delivering not only increased security but also more efficiency and agility in the defence against cyber attacks in a world that is becoming increasingly digital.

Keyword: Artificial Intelligence, machine learning, cyber security, technologies, cyber-attacks.

INTRODUCTION

Because this article is about knowing how Artificial Intelligence (AI) is going to be used and how it will be used in cyber security duties, it is of the utmost essential to have a solid knowledge of what AI is. The goal of artificial intelligence is to construct and identify intelligent objects. According to Russell and Norvig (2016), artificial intelligence is a wide field since it allows users or employees of the technology to apply their abilities to any sector that they feel appropriate. When the machine begins to imitate human intellect and begins to learn on its own, it leads to solutions that were previously unknown to the machine. This is something that the machine was never capable of achieving before. Algorithmic modelling, on the other hand, has made this a reality, and artificial intelligence has become a standard word and function. Furthermore, artificial intelligence technologies such as expert systems, machine learning, deep learning or neural networks, artificial immune systems, intelligent agents, and so on are being utilised in a variety of disciplines, including but not limited to the healthcare industry, the automobile industry, the banking industry, and the insurance industry.

A significant challenge that organisations face in the modern era is figuring out how to protect themselves from the possibility of anomalies. Both the diversity and the probability of these unexplained breakouts create the necessity for corporations to place a high priority on the manner by which they protect themselves against cyber attacks of this kind. Based on the findings of Felder, Panaousis, Malacaria, Hankin, and Smeraldi (2016), it is important for any corporation or firm to have a comprehensive understanding of the assaults that they are most vulnerable to in order to minimise the likelihood of being attacked. The organisation as a whole is seeing a significant increase in the level of uncertainty as a direct result of the recent surge in the number of cyber attacks.

Due to the fact that cyber attacks on global enterprise networks, including those of governments, have kept cyber attack prevention teams exceedingly busy, automation alone will not be sufficient to meet the demands of the situation. When taking into consideration the development of artificial intelligence capabilities or techniques, such as neural networks or deep learning, machine learning, expert systems, artificial immune systems, and intelligent agents, there is a growing demand for proactive measures to be taken in response to cyber attacks of this nature. These measures should include the ability to anticipate those attacks and find solutions to them. The management of cyber security in modern businesses, which are more susceptible to assaults that are unknown or unforeseen than they have ever been before, is being altered by the use of such cutting-edge technology such as these. It is believed that the transformation will not only involve the proactive management of cyber attacks, but also the prediction of such assaults and the resolution of issues before they have a negative impact on the operations of a company and put the data of its customers at danger.

The term "artificial intelligence" (AI) refers to the capacity of a digital computer or a robot controlled by a computer to carry out activities that are traditional associated with intelligent individuals. In the context of the endeavour of constructing systems that are endowed with the intellectual processes that are distinctive of humans, such as the ability to reason, uncover meaning, generalise, or learn from previous experience, the phrase is commonly utilised. In the decades after their invention in the 1940s, digital computers have been programmed to perform extremely difficult jobs with a high level of expertise. These activities include the discovery of proofs for mathematical theorems and the ability to play chess. Although there have been ongoing advancements in the speed of computer processing and the capacity of memory, there are still no programs that can match the complete human flexibility in a wider range of fields or in jobs that need a significant amount of everyday knowledge. On the other hand, some programs have attained the performance levels of human experts and professionals in executing certain specific tasks, so that artificial intelligence in this limited sense is found in applications as diverse as medical diagnosis, computer search engines, voice or handwriting recognition, and chatbots.

The primary purpose of cyber security is to safeguard the electronic devices that we all use, such as smartphones,

laptops, tablets, and computers, as well as the services that we use, both online and at our places of employment, from being stolen or damaged.

Additionally, it is about avoiding unauthorised access to the massive quantities of personal information that we save on these devices and online. There is a significant need for cyber security since mobile devices, computers, and the internet have become such an integral component of contemporary life that it is impossible to conceive of how we might operate in the absence of these technologies. It is more vital than ever before to adopt measures that can prevent cybercriminals from gaining access to our accounts, data, and devices. These measures can be taken in connection with online banking and shopping, as well as email and social media.

OBJECTIVE

1. To understand causes of the knowledge acquisition problem impact on expert systems in decision-making or problem-solving to improve the management of cyberattacks.

2. To explore how cyberattacks can be prevented proactively using AI neural networks to prevent malicious unknown intrusions.

CORE APPROACHES TO ARTIFICIAL INTELLIGENCE TECHNOLOGY IN THREAT DETECTION

The application of artificial intelligence technology in the field of cyber security threat detection significantly enhances the accuracy and efficiency of detection by virtue of its powerful data analysis and pattern recognition capabilities. Among the key methodologies, machine learning, deep learning, integrated learning, and multimodal methods are the most prominent examples. Integrated learning and multimodal approaches further improve the robustness and adaptability of detection by combining multiple data sources and algorithms. Machine learning is able to learn threat features from large amounts of labelled data, deep learning processes complex unstructured data through multi-layer neural networks, and integrated learning and multimodal approaches are able to learn threat features from large amounts of labelled data. Below, we will go into further depth on these three fundamental techniques. The Cross-Entropy Loss, also known as the Binary Classification Loss:

$$\text{Loss} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$$

Machine Learning in Threat Detection

Training models to recognise known threat categories in fresh data is accomplished through the use of supervised learning, which is one of the most prevalent methodologies in machine learning. This method is dependent on labelled datasets. The field of cyber security makes extensive use of supervised learning in a variety of contexts, including the identification of malware, the filtering of spam, and the detection of intrusions. A significant quantity of data that has been labelled as normal and anomalous enables the model to acquire the ability to learn and recognise the features of harmful behaviours. As a result, it is able to successfully differentiate between regular traffic and threat activities in real-time detection applications. The high accuracy and interpretability of this technique are two of its notable benefits; nevertheless, the efficacy of this strategy is contingent on the quality and amount of labelled data, and it may have some limits when it comes to threats that are new or unknown.

Unsupervised learning does not rely on labelled datasets; rather, it examines the inherent structure of the data in order to uncover hidden patterns and behaviours that are not typical of the data. Anomaly detection and intrusion detection are the two primary applications of unsupervised learning in the field of cybersecurity. This sort of

learning is particularly well-suited for identifying threat categories that are not conventionally characterised. The techniques of cluster analysis and anomaly detection, such as K-means, isolated forests, and self encoders, are examples of common unsupervised learning approaches. These techniques are able to recognise unusual behaviours that deviate from typical patterns of behaviour, and as a result, they may uncover possible threats to security. The capacity of unsupervised learning to identify unknown dangers is one of its advantages; nevertheless, it also has a high risk of false positives, which means that it must be optimised in conjunction with other methods in order to get optimal results.

The machine learning technique known as reinforcement learning is a dynamic decision-making approach that is appropriate for situations that call for ongoing learning and adaptability to changes in the surrounding environment. When it comes to the detection of threats to network security, reinforcement learning is used to make constant adjustments to the detection method in order to maximise the effectiveness of security protection by interacting with the environment. Reward learning may be utilised, for instance, to dynamically alter the alarm levels of an intrusion detection system or to optimise the defence strategy in response to changing assault strategies. Both of these applications are examples of how reinforcement learning can be utilised. As shown in Figure 1, reinforcement learning is more flexible in comparison to older approaches. It is also capable of updating detection algorithms in real time in accordance with new threat data, which improves both the flexibility and the real-time network security protection:

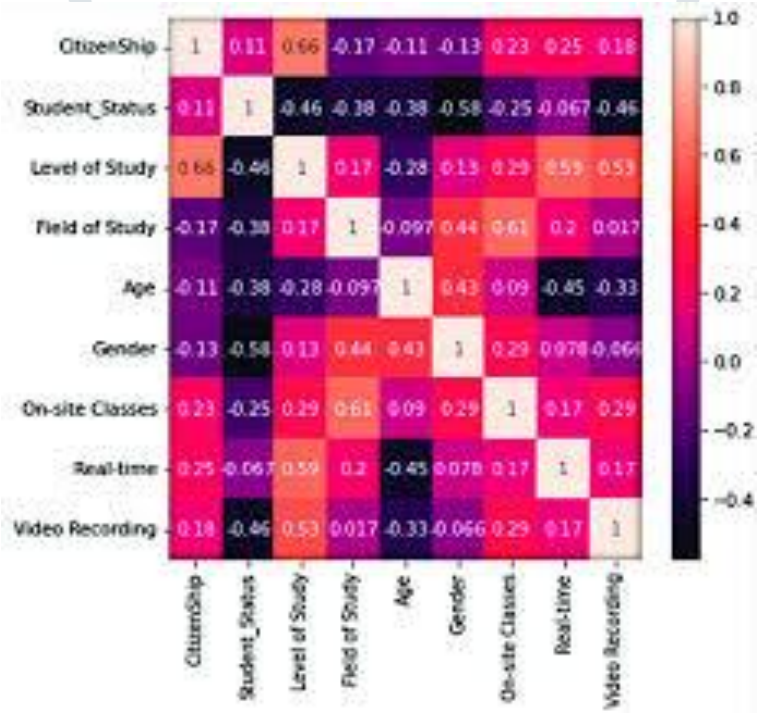

Figure 1. Correlation Matrix of Features

Transfer learning is a method that makes use of previously acquired information in order to successfully complete new tasks. This method is particularly well-suited for use in circumstances in which there is a limited amount of data or where the costs of training are high. When it comes to cyber security, migration learning has the potential to significantly enhance the efficiency of detection models. This is accomplished by transferring threat detection information acquired from one domain (for example, financial security) to another domain (for example, healthcare security). In addition to minimising the amount of time needed for model training and the amount of computational resources that are used, migration learning makes it possible to construct effective threat detection models with minimal data. However, the difficulty of migration learning comes in the fact that it is necessary to pick the degree of similarity between the source domain and the destination domain with great

care in order to guarantee that the knowledge migrating process is successful. A Rule for the Gradient Descent Update:

$$\theta_{t+1} = \theta_t - \alpha \nabla_\theta J(\theta) \quad (2)$$

Deep Learning in Threat Detection

Given that deep learning is able to automatically extract and learn aspects of complicated data through the design of multi-layer neural networks, it has a huge edge when it comes to the identification of cybersecurity threats. Convolutional Neural Networks, often known as CNNs, are frequently employed for the purpose of processing picture and traffic data. These networks are able to successfully identify characteristics such as malicious code or abnormal network traffic. On the other hand, Recurrent Neural Networks (RNN) and its upgraded version, Long Short-Term Memory Networks (LSTM), are effective at processing time-series data and are suited for application situations such as analyzing network behaviour logs and identifying persistent threats. Both of these networks are able to process time-series data without any problems. The ability of these neural network models to identify threats that are difficult to detect using standard approaches is improved by deep learning of large-scale data. This results in an increase in both the accuracy of detection and the speed with which responses are made.

Through the process of compressing and rebuilding the input data, an auto-encoder is a type of unsupervised learning model that is capable of learning a latent representation of the data. Auto-encoders are frequently utilised in the field of cyber security for the purpose of anomaly detection. This includes the identification of aberrant network traffic or patterns of user behaviour. There is a large rise in the reconstruction error if the input data is significantly different from the regular data patterns. This phenomenon indicates the presence of potential dangers. On the other hand, Generative Adversarial Networks (GANs) are made up of a generator and a discriminator. Their purpose is to generate false data that is similar to actual data and to enhance the capability of the detection model through adversarial training. In the field of cyber security, GANs may be utilised not only for the purpose of generating attack samples in order to increase the resilience of the detection model, but also for the purpose of spoofing the detection system in order to carry out adversarial tests in order to assess and improve the system's capacity to provide protection.

In contrast to deep learning models, which often need a substantial quantity of data and computer resources for training, migration learning may rapidly construct efficient threat detection models with a smaller amount of data by making use of models that have already been extensively trained. Pre-trained models are often trained on large-scale datasets, and then they are applied to particular cyber security tasks such as the categorisation of malware or the detection of aberrant traffic. Migration learning allows deep learning models to inherit the knowledge of pre-trained models, swiftly adapt to new threat scenarios, and enhance detection outcomes. This is made possible through the process of migration learning. This strategy not only reduces the amount of time and resources required for training, but it also enhances the generalisation capability of the model, which enables it to deal with a wide variety of security risks on its own.
A Graph Neural Network, often known as a GNN, is a deep learning model that specialises in the processing of graph-structured data. This model is particularly well-suited for network topology analysis and threat identification. When it comes to the realm of cybersecurity, GNNs have the capability to process complex structured data, such as social networks and computer networks, in order to identify possible dangers. This is accomplished by capturing the links that exist between nodes. As an illustration, GNN may be utilised to identify unusual communication patterns inside a network, as well as to uncover hidden hostile nodes or attack vectors

concealed within the network. When compared to more conventional approaches, GNN is able to better comprehend the connections that exist inside intricate network architectures. Furthermore, it is able to deliver more precise findings for threat identification, hence delivering novel concepts and instruments for the protection of network.

Integrated Learning and Multimodal Approaches

Through the process of merging the outcomes of many models' predictions, integrated learning enables an improvement in the overall detection accuracy and resilience. When it comes to the detection of cybersecurity risks, integrated learning has the ability to combine the benefits of many algorithms. For example, it may integrate several models, such as decision trees, random forests, gradient boosting trees, and so on, in order to enhance the identification of complicated threats. Through the complementarity of many models, the integrated learning approach has the potential to effectively minimise the false alarm rate of a single model and boost the flexibility to a wide range of threats. When it comes to the detection of malware, for instance, integrated learning may be used to combine a number of different feature extraction algorithms and classifiers in order to enhance the detection process's accuracy and capacity to generalise.

Data that is engaged in the process of detecting threats to network security is often multimodal. This data may include data on network traffic, system logs, user behaviour, geolocation, and other similar information. Through the combination of information from a variety of data sources, multimodal techniques make it possible to gain a more thorough picture of potential dangers and to identify them. By integrating data on network traffic with data on user behaviour, for instance, it is possible to spot aberrant behaviour and probable attack patterns with greater precision. It is possible for multimodal data fusion to increase the perceptual capacity and decision-making accuracy of the detection system. Additionally, it may develop a multidimensional knowledge of complex threats by completely analysing input from several modalities, which ultimately results in an improvement in the overall effect of threat detection.

A considerable improvement in the efficiency of threat detection may be achieved by the utilisation of heterogeneous integration methods, which are an extension of integration learning. These approaches mix many types of models and several distinct data sources. Heterogeneous integration methods have more processing power than traditional isomorphic integration methods (for example, the integration of multiple models that are similar to one another). This is because they integrate the benefits of different models (for example, deep learning models and traditional machine learning models), which is especially beneficial when dealing with unknown threats and complex attacks. For instance, this heterogeneous integration strategy can fully utilise the powerful feature representation capability of deep learning and the decision making capability of integrated learning to achieve more accurate threat detection, as shown in Figure 2. This is accomplished by using deep learning models for feature extraction and feeding their outputs into integrated learning models for final decision making during the process:

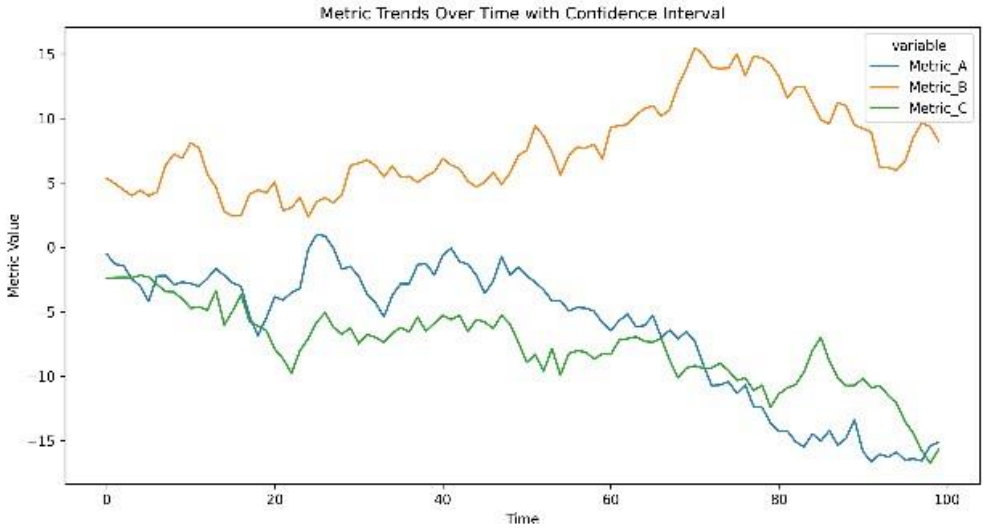

Figure 2. Metric Trends Over Time with Confidence Interval

Through the simultaneous analysis of various data modalities, multimodal deep learning may more effectively address complex cybersecurity threats. When applied in real-world scenarios, multimodal deep learning models have the ability to combine several forms of data, including text, photos, audio, and video, among others, in order to conduct complete threat analysis and detection. As an illustration, multimodal deep learning may be utilised in the detection of APT (Advanced Persistent Threat) by combining network data, log analysis, and user behaviour in order to offer a global perspective and identify dangers that have been hiding for a long time. Multimodal deep learning has the benefit of being able to combine different information sources in order to generate a holistic judgement of the threat. This helps to increase the accuracy and robustness of detection, particularly when dealing with sophisticated threats and data that is multi-dimensional.

4. CHALLENGES AND FUTURE DEVELOPMENTS OF ARTIFICIAL INTELLIGENCE IN CYBERSECURITY THREAT DETECTION

The most significant obstacles that arise throughout the process of using AI for the identification of cyber security threats are those pertaining to data privacy and security. It is possible that sensitive information is contained among this data since threat detection is dependent on a huge quantity of data, which includes network traffic, user behaviour, system logs, and other similar data. During the process of model training and deployment, there is an essential challenge that has to be handled, and that is how to safeguard these data from being misused or leaked. Currently, techniques that protect privacy, such as differential privacy and federated learning, are gradually being applied to the field of cyber security. However, the question of how to guarantee the accuracy and efficiency of the detection model while simultaneously protecting data privacy is still a direction that needs to be further investigated.

The identification of threats is an area in which artificial intelligence models thrive, but they also confront the challenge of being attacked by adversaries. By creating adversarial samples or in some other way, an adversary might fool the model into making incorrect judgment's, which ultimately results in the security system failing to function properly. These types of adversarial assaults not only reduce the efficiency of the detection system, but they also have the potential to be utilised in order to successfully defeat security defences. Therefore, increasing the resilience of artificial intelligence models. Increasing the model's resilience to assaults, constructing architectures that are more secure, and implementing protective mechanisms are all ways that researchers are investigating how to increase the dependability of threat detection systems. Autoencoder-Based Anomaly Detection and Analysis (Reconstruction Error):

$$\text{Reconstruction Error} = \frac{1}{N}\sum_{i=1}^{N}|x_i - \hat{x}_i|^2 \quad (3)$$

The decision-making processes of artificial intelligence models, particularly deep learning models, are sometimes referred to as "black boxes" since they are difficult to explain due to their complexity. Model interpretability is of the utmost importance in the field of cybersecurity, as the findings of threat detection are frequently need to be confirmed and evaluated by security professionals. In addition, compliance standards are growing more demanding as an increasing number of companies and domains continue to embrace artificial intelligence for the purpose of protecting their facilities. An important problem that will need to be solved in the future is figuring out how to increase the performance of models while simultaneously improving their interpretability in order to comply with a variety of different sorts of privacy and security standards. Explainable Artificial Intelligence (XAI) and model visualisation tools are being increasingly implemented in order to assist security professionals in comprehending and having faith in the outcomes of detection.

In the future, as technology continues to improve, threat detection systems will become more intelligent and automated, and they will be able to respond to new and sophisticated threats in a more timely and accurate manner. The application of artificial intelligence in cyber security threat detection continues to hold a great deal of promise. Cooperation between different fields and disciplines will make it easier for more inventive applications to come into existence. One example of this would be the development of threat detection systems that integrate artificial intelligence with digital block chain technology and quantum computing. In addition, the application of artificial intelligence in cyber security will become safer, more reliable, and more efficient as a result of the ongoing development of strategies for counter-attack protection, mechanisms for privacy protection, and interpretable models. In the not too distant future, artificial intelligence will not only serve as a tool for threat detection, but it will also become an essential component in cyber security strategy, therefore driving the whole sector towards a safer level of protection.

CONCLUSION

The use and usage of artificial intelligence and cyber security are complimentary to one another in terms of their respective applications. When it comes to deploying expert systems to defend against cyber attacks, the most important aspect is the accumulation of knowledge. Bringing about a significant beneficial change in the way risks are now managed in order to enhance this vulnerable area may be accomplished through the use of quality, completeness, and accuracy. When it comes to the construction of expert systems, a good option is one that is founded on accurate knowledge representation. In the context of applications involving expert systems, the selection of a representation approach to begin with is of particular and paramount significance. It is common knowledge that neural networks, which include both shallow learning and deep learning, have advanced the cyber security solution to a higher level than it was previously previously. These networks have evolved into cutting-edge technology that are used to combat cyber attacks. The only data that is supplied to deep learning neural networks is the input data, which allows them to discover patterns from the data on their own autonomy. When it comes to this type of learning that is not supervised, the data that is to be found is referred to as unlabelled data and is categorised as Self Organisation Maps (SOMs) and Adaptive Resonance Theory (ART). Among the several intrusion detection systems, the hybrid artificial neural network approach is the best suitable option in terms of detection rates, false positives, false negatives, as well as cost and time savings. To combat distributed denial of service attacks (DDoS), it is likely that intelligent agents that are able to communicate and comprehend language will take on the role of cyber police. Distributed multi-agent systems are able to learn on their own from past anomalies and are able to learn from history. Their design philosophy and functional benefits, such as being autonomous and collaborating, make it possible to produce real-time and distributed DDoS discovery or detection systems, as well as systems that are able to simultaneously find and shut down a variety of sources or attacks. In the event of a distributed denial of service attack, it is possible to dispatch

numerous mobile agents to the users who are affected in order to investigate and mitigate the effects of network abnormalities.

REFERENCES

1. Anwar, A., & Hassan, S. I. (2017). Applying artificial intelligence techniques to prevent cyber assaults. International Journal of Computational Intelligence Research, 13(5), 883–889

2. Calderon, R. (2019). The Benefits of Artificial Intelligence in Cybersecurity. Economic Crime Forensics Capstones. 36. https://digitalcommons.lasalle.edu/ecf_capstones/36

3. G. Aceto, D. Ciuonzo, A. Montieri, and A. Pescapé, "Distiller: Encrypted traffic classification via Multimodal Multitask Deep Learning," Journal of Network and Computer Applications, vol. 183-184, p. 102985, 2021

4. C. Islam, M. A. Babar, R. Croft, and H. Janicke, "SmartValidator: A framework for automatic identification and classification of cyber threat data," Journal of Network and Computer Applications, vol. 202, p. 103370, Jun. 2022, doi: 10.1016/j.jnca.2022.103370.

5. Xu S, Qian Y, Hu R Q. Data-Driven Network Intelligence for Anomaly Detection [J]. IEEE Network, 2019, 33(3):88-95. DOI:10.1109/MNET.2019.1800358.

6. Silvestri S, Islam S, Amelin S C M. Cyber threat assessment and management for securing healthcare ecosystems using natural language processing [J]. International Journal of Information Security, 2024, 23(1):31-50

7. Iwendi C, Rehman S U, Javed A R, et al. Sustainable Security for the Internet of Things Using Artificial Intelligence Architectures [J]. ACM Transactions on Internet Technology, 2021, 21(3):1-22. DOI:10.1145/3448614.

8.] Zhao L, Zhu D, Shafik W, et al. Artificial intelligence analysis in cyber domain: A review: [J]. International Journal of Distributed Sensor Networks, 2022, 18(4):121-131. DOI:10.1177/15501329221084882.

9. Wikipedia Contributors, "Cyber security," Wikipedia, Apr. 29, 2019. [Online]. Available: https://en.wikipedia.org/wiki/Machine_learning

10. Nicolau, A. D. S., Augusto, J. P. D. S., & Schirru, R. (2017, June). Accident diagnosis system based on realtime decision tree expert system. In AIP Conference Proceedings (Vol. 1836, No. 1, p. 020017). AIP Publishing LLC. doi:10.1063/1.4981957

11. Patil, P. (2016). Artificial intelligence in cybersecurity. International Journal of Research in Computer Applications and Robotics, 4(5), 1–5

12. Wirkuttis, N., & Klein, H. (2017). Artificial intelligence in cybersecurity. Cyber Intelligence and Security Journal, 1(1), 21–23.

13. Shankar, S. (2017). Looking into the Black Box. Holding Intelligent Agents Accountable. NUJS L. Rev., 10, 451.

14. Dutt, I., Borah, S., & Maitra, I. (2016). Intrusion detection system using artificial immune system. International Journal of Computers and Applications, 144(12)