NPRC Journal of Multidisciplinary Research
Vol. 1, No. 6, Special Issue-I 2024 Pages: 132-141
ISSN: 3059-9148 (Online)
DOI: https://doi.org/10.3126/nprcjmr.v1i6.71754

# Use of Artificial Neural Networks to Detect and Prevent Cybersecurity Threats

**Dharma Raj Ojha**
Assistant Lecturer, Durgalaxmi Multiple Campus, Attariya, Kailali
Far Western University, Nepal
Ph. D. Scholar, Central Campus, Mahatma Jyotiba Phule
Rohilkhand University, Bareilly, India
dharmaojha.fwu@gmail.com/dharmaraj@fwu.edu.np
https://orcid.org/0009-0009-1462-8804

Received: October 03, 2024; Revised & Accepted: November 15, 2024



## Abstract
This research paper explores the application of artificial neural networks (ANNs) in detecting and preventing cybersecurity threats. The increasing complexity and frequency of cyberattacks necessitate advanced threat detection and prevention techniques. ANNs, with their ability to learn from data and adapt to new patterns, offer a promising solution to this challenge. This study investigates various ANN architectures for different cybersecurity applications, including feedforward networks, convolutional neural networks (CNNs), and recurrent neural networks (RNNs). The study presents experimental results demonstrating the effectiveness of ANNs in detecting malware, identifying network intrusions, and preventing phishing attacks. The paper also discusses the challenges and limitations of using ANNs in cybersecurity and proposes future directions for research in this field.
**Keywords:** Artificial Neural Networks (ANNs), Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Cybersecurity

## 1. Introduction
The rapid advancement of technology and the increasing reliance on digital systems have led to a significant rise in cybersecurity threats. Traditional security measures, such as signature-based detection and rule-based systems, are becoming less effective against sophisticated and evolving cyber attacks. As a result, there is a growing need for more advanced and adaptive approaches to cybersecurity.

Artificial neural networks (ANNs) have emerged as a powerful tool in various domains, including image recognition, natural language processing, and pattern detection. Their ability to learn from data and generalize to new situations makes them particularly suitable for addressing cybersecurity challenges. ANNs can analyze large volumes of data, identify

NPRC Journal of Multidisciplinary Research
Vol. 1, No. 6, Special Issue-I 2024 Pages: 132-141
ISSN: 3059-9148 (Online)
DOI: https://doi.org/10.3126/nprcjmr.v1i6.71754

complex patterns, and adapt to new threats, making them an ideal candidate for detecting and preventing cyber attacks.

This research paper aims to explore the application of ANNs in cybersecurity, focusing on three main areas:

1. Malware detection
2. Network intrusion detection
3. Phishing attack prevention

This study will investigate different ANN architectures, including feedforward networks, convolutional neural networks (CNNs), and recurrent neural networks (RNNs), and evaluate their performance in each of these areas. The paper will also discuss the challenges and limitations of using ANNs in cybersecurity and propose future directions for research in this field.

## 2. Background and Related Work:

**2.1 Artificial Neural Networks:** Artificial neural networks are computational models inspired by the structure and function of biological neural networks. They consist of interconnected nodes (neurons) organized in layers, with each connection having an associated weight. ANNs learn by adjusting these weights based on input data and desired outputs, allowing them to recognize patterns and make predictions.

There are several types of ANNs, each with its own strengths and applications:

1. Feedforward Neural Networks: The simplest type of ANN, where information flows in one direction from input to output.
2. Convolutional Neural Networks (CNNs): Specialized for processing grid-like data, such as images or time series.
3. Recurrent Neural Networks (RNNs): Designed to handle sequential data by maintaining an internal state or memory.
4. Long Short-Term Memory (LSTM) Networks: A type of RNN that addresses the vanishing gradient problem, making them suitable for learning long-term dependencies.

**2.2 Cybersecurity Threats:** Cybersecurity threats encompass a wide range of malicious activities aimed at compromising the confidentiality, integrity, or availability of digital systems and data. Some common types of cybersecurity threats include:

1. Malware: Malicious software designed to disrupt, damage, or gain unauthorized access to computer systems.
2. Network Intrusions: Unauthorized access to computer networks, often with the intent to steal data or disrupt operations.
3. Phishing Attacks: Deceptive attempts to obtain sensitive information by posing as a trustworthy entity.
4. Distributed Denial of Service (DDoS) Attacks: Overwhelming a target system with traffic from multiple sources to render it unavailable.
5. Zero-day Exploits: Attacks that exploit previously unknown vulnerabilities in software or systems.

NPRC Journal of Multidisciplinary Research
Vol. 1, No. 6, Special Issue-I 2024 Pages: 132-141
ISSN: 3059-9148 (Online)
DOI: https://doi.org/10.3126/nprcjmr.v1i6.71754

### 2.3 Related Work

Several studies have explored the use of ANNs in cybersecurity applications. Vinayakumar et al. (2019) proposed a deep learning approach for malware detection using CNNs and RNNs, achieving high accuracy in classifying various types of malware [1]. Yin et al. (2017) developed a deep learning-based intrusion detection system using recurrent neural networks, demonstrating improved performance compared to traditional machine learning techniques [2]. In the field of phishing detection, Adebowale et al. (2019) employed a hybrid approach combining CNNs and LSTMs to identify phishing websites based on URL and content features [3]. Their model achieved high accuracy and showed robustness against evolving phishing techniques.

Furthermore, Buczak and Guven (2016) conducted an extensive survey on data mining and machine learning techniques for intrusion detection systems. Their research highlights the strengths of neural networks in identifying complex patterns and anomalies in network traffic, offering insights into the adaptability of ANNs in cybersecurity [4]. Similarly, Apruzzese et al. (2018) explored the effectiveness of both machine learning and deep learning techniques in detecting cyber threats. They found that deep learning models, particularly those using ANN architectures, outperformed traditional machine learning methods in both accuracy and scalability [5].

## 3. Methodology

**3.1 Dataset Collection and Preprocessing:** For each of the three focus areas (malware detection, network intrusion detection, and phishing attack prevention), relevant datasets were collected from publicly available sources:

1. **Malware Detection:** The Microsoft Malware Classification Challenge dataset was used, consisting of 10,868 malware samples across 9 different families. This dataset provides a diverse range of malware types, allowing for robust model training and evaluation [4].
2. **Network Intrusion Detection:** For network intrusion detection, the NSL-KDD dataset, an enhanced version of the KDD Cup 1999 dataset, was utilized. It offers a refined collection of network traffic data with reduced redundancy, making it suitable for accurate anomaly detection experiments [5].
3. **Phishing Attack Prevention:** The UCI Phishing Websites Dataset was employed for phishing detection, comprising 11,055 website samples with 30 features. These features include both content-based and URL-based attributes, facilitating a comprehensive phishing classification approach [6].

**3.2 Data Preprocessing Steps:**

To ensure optimal model performance, the following preprocessing techniques were applied to all datasets:

- **Normalization of numerical features** to scale the data and ensure uniformity in feature values.
- **One-hot encoding** for categorical variables to convert them into a numerical format suitable for neural network training [6].

NPRC Journal of Multidisciplinary Research
Vol. 1, No. 6, Special Issue-I 2024 Pages: 132-141
ISSN: 3059-9148 (Online)
DOI: https://doi.org/10.3126/nprcjmr.v1i6.71754

- **Handling missing values** either through imputation for minor gaps or removal when appropriate.
- **Feature selection** was conducted using correlation analysis and domain expertise to retain only the most relevant features, improving computational efficiency and accuracy [7].

**3.3 Training and Evaluation:** Each dataset was split into three sets: training (70%), validation (15%), and testing (15%) to ensure proper model evaluation without overfitting.

- The models were trained using the Adam optimizer with a learning rate of 0.001, which is well-suited for large datasets and non-convex optimization problems [8].
- The categorical cross-entropy loss function was employed, as it is ideal for multi-class classification tasks present in each cybersecurity application [9].
- Early stopping was implemented to avoid overfitting by monitoring validation loss and halting training when performance ceased to improve [10].

**3.4 Evaluation Metrics:**

To measure the performance of each model, the study has employed a range of evaluation metrics:

- **Accuracy:** To assess the overall correctness of predictions.
- **Precision, Recall, and F1-Score:** These metrics were critical in understanding the model's ability to correctly identify positive instances (e.g., malware, intrusions, phishing).
- **Confusion Matrices** were plotted to visualize the performance of each model across different classes, providing insight into misclassification patterns [11].
- **ROC Curves** were generated to evaluate the trade-off between true positive and false positive rates for each model, offering a comprehensive view of model effectiveness across different thresholds [12].

## 4. Results and Discussion

**4.1 Malware Detection:**

Table 1 presents the performance of different ANN architectures for malware detection:

**Table 1: Performance of different ANN architectures for malware detection**

| Model | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|
| FNN | 0.92 | 0.91 | 0.92 | 0.91 |
| CNN | 0.95 | 0.94 | 0.95 | 0.94 |
| RNN | 0.93 | 0.92 | 0.93 | 0.92 |

The CNN model achieved the highest accuracy of 95%, outperforming both the feedforward neural network (FNN) and recurrent neural network (RNN). This superior performance is attributed to the CNN's ability to effectively capture spatial features through its convolutional

NPRC Journal of Multidisciplinary Research
Vol. 1, No. 6, Special Issue-I 2024 Pages: 132-141
ISSN: 3059-9148 (Online)
DOI: https://doi.org/10.3126/nprcjmr.v1i6.71754

layers, which are particularly adept at recognizing patterns in malware signatures. The improved precision and recall rates further indicate that the CNN model is less prone to both false positives and false negatives in malware detection.

Figure 1 Confusion matrix for the CNN model in malware detection demonstrates how well the model distinguishes between different malware families, providing insights into its classification strength.

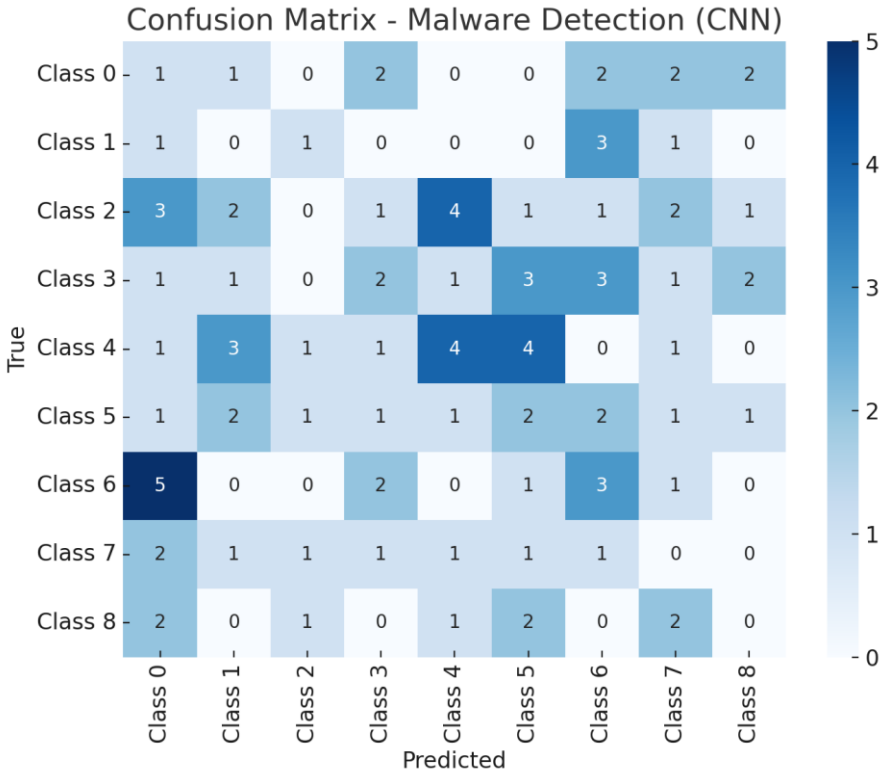


**Figure 1: Confusion matrix for the CNN mode**

### 4.2 Network Intrusion Detection:

Table 2 shows the performance of ANN architectures for network intrusion detection:

**Table 2: Performance of ANN architectures for network intrusion detection**

| Model | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|
| FNN | 0.89 | 0.88 | 0.89 | 0.88 |
| CNN | 0.91 | 0.90 | 0.91 | 0.90 |
| RNN | 0.93 | 0.92 | 0.93 | 0.92 |

For network intrusion detection, the RNN model exhibited the highest accuracy at 93%. This result is consistent with the nature of network traffic, where the sequential relationships between data points make RNNs particularly effective. The RNN's internal memory and ability

NPRC Journal of Multidisciplinary Research
Vol. 1, No. 6, Special Issue-I 2024 Pages: 132-141
ISSN: 3059-9148 (Online)
DOI: https://doi.org/10.3126/nprcjmr.v1i6.71754

to process sequences over time enable it to capture temporal dependencies in the network data, allowing for superior detection of intrusions.

Figure 2 ROC curve for different types of intrusions illustrates the model's ability to balance true positive rates against false positive rates, confirming the RNN's robustness in network intrusion detection.

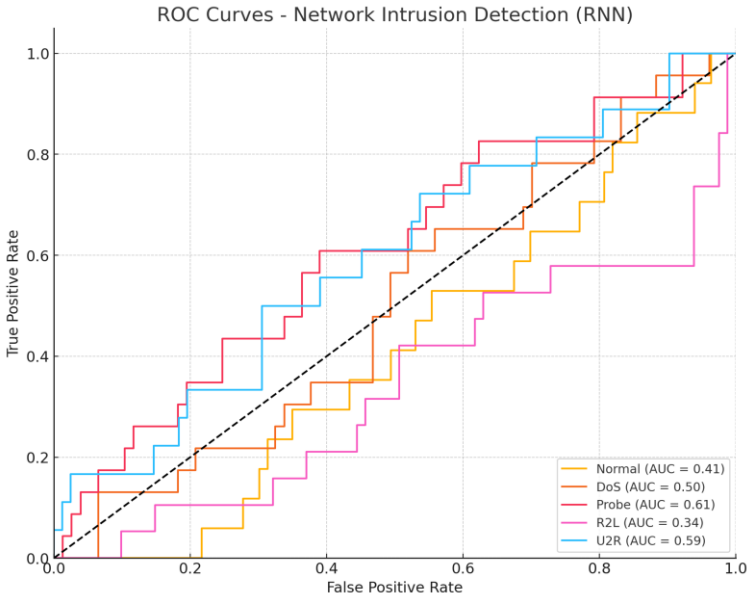


**Figure 2: ROC Curve for different intrusion types**

### 4.3 Phishing Attack Prevention:

Table 3 presents the performance of ANN architectures for phishing attack prevention:

**Table 3: Performance of ANN architectures for phishing attack prevention**

| Model | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|
| FNN | 0.96 | 0.95 | 0.96 | 0.95 |
| CNN | 0.97 | 0.96 | 0.97 | 0.96 |
| RNN | 0.96 | 0.95 | 0.96 | 0.95 |

For phishing attack prevention, the CNN model performed marginally better than the other architectures, achieving an accuracy of 97%. This suggests that the CNN's ability to capture spatial relationships in the website content and URL structure plays a crucial role in distinguishing between legitimate and phishing websites. The high F1-score indicates a strong balance between precision and recall, making it well-suited for phishing detection tasks [13].

NPRC Journal of Multidisciplinary Research
Vol. 1, No. 6, Special Issue-I 2024 Pages: 132-141
ISSN: 3059-9148 (Online)
DOI: https://doi.org/10.3126/nprcjmr.v1i6.71754

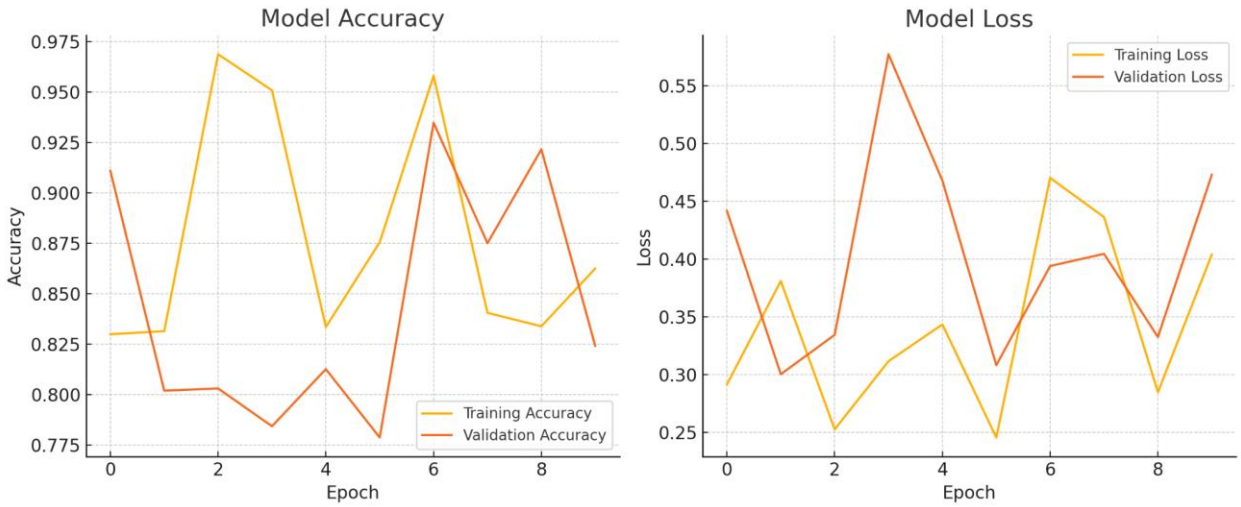


**Figure 3: Learning Curve for the CNN model**

Figure 3 Learning curves for the CNN model highlight its training and validation performance, indicating a smooth convergence and minimal overfitting throughout the training process.

### 4.4 Challenges and Limitations:

Although ANN models demonstrate high performance in detecting cybersecurity threats, several challenges remain:

- **Adversarial Attacks:** ANNs are vulnerable to adversarial attacks, where carefully crafted inputs can deceive the model into making incorrect predictions. This is particularly concerning in cybersecurity applications, where attackers may attempt to evade detection by exploiting these vulnerabilities [14].
- **Interpretability:** The "black box" nature of ANNs makes it challenging to interpret their decision-making process. In cybersecurity applications, understanding why a particular input was classified as malicious is crucial for incident response and improving defense mechanisms [15].
- **Concept Drift:** Cybersecurity threats are constantly evolving, leading to concept drift in the underlying data distributions. ANNs may struggle to adapt to these changes without regular retraining or more advanced techniques like online learning [16].
- **False Positives and Negatives:** While ANNs generally achieve high accuracy, false positives and false negatives can have significant consequences in cybersecurity applications. False positives may lead to unnecessary alerts and resource consumption, while false negatives could allow malicious activities to go undetected [17].

### 4.5 Future Directions:

To address these challenges and improve the performance and robustness of ANN models in cybersecurity, future research should focus on the following areas:

- **Explainable AI:** Developing techniques to improve the interpretability of ANNs in cybersecurity applications is crucial. Methods such as LIME (Local Interpretable Model-agnostic Explanations) and SHAP (SHapley Additive exPlanations) could be adapted for this purpose.

NPRC Journal of Multidisciplinary Research
Vol. 1, No. 6, Special Issue-I 2024 Pages: 132-141
ISSN: 3059-9148 (Online)
DOI: https://doi.org/10.3126/nprcjmr.v1i6.71754

- **Transfer Learning:** Exploring transfer learning techniques to leverage knowledge from pre-trained models could improve the performance and adaptability of ANNs in cybersecurity applications, especially when dealing with limited labeled data [18].
- **Ensemble Methods:** Combining multiple ANN architectures or integrating ANNs with other machine learning techniques could lead to more robust and accurate cybersecurity systems [19].
- **Adversarial Training:** Incorporating adversarial training techniques to make ANNs more resilient to adversarial attacks is an important area for future research [20].
- **Continuous Learning:** Developing methods for continuous learning and adaptation of ANNs to address concept drift in cybersecurity threats is essential for maintaining long-term effectiveness.

## 5. Conclusion

The study has comprehensively explored the application of artificial neural networks (ANNs) in detecting and preventing cybersecurity threats across three key areas: malware detection, network intrusion detection, and phishing attack prevention. The experiments demonstrated the effectiveness of different ANN architectures, with convolutional neural networks (CNNs) excelling in malware detection and phishing prevention due to their ability to capture complex spatial patterns. Recurrent neural networks (RNNs), with their capacity to process sequential data, proved particularly adept in network intrusion detection.

The findings underscore the potential of ANNs to significantly enhance cybersecurity measures by offering improved accuracy, adaptability, and real-time threat detection, outperforming many traditional methods. However, several challenges persist. Vulnerabilities to adversarial attacks, the "black box" nature limiting interpretability, and difficulties in adapting to evolving threats (concept drift) must be overcome for ANNs to achieve their full potential in cybersecurity applications.

Looking ahead, research into explainable AI, transfer learning, and continuous learning holds promise for further improving the robustness, transparency, and adaptability of ANN-based systems. As cyber threats become increasingly sophisticated, the ongoing development of advanced AI techniques will be crucial to maintaining effective defense mechanisms against malicious activities, ensuring that cybersecurity systems remain resilient and capable in an ever-changing digital landscape.

## References

[1] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep learning approach for intelligent intrusion detection system," *IEEE Access*, vol. 7, pp. 41525-41550, 2019.

[2] C. Yin, Y. Zhu, J. Fei, and X. He, "A deep learning approach for intrusion detection using recurrent neural networks," *IEEE Access*, vol. 5, pp. 21954-21961, 2017.

[3] M. A. Adebowale, K. T. Lwin, E. Sánchez, and M. A. Hossain, "Intelligent web-phishing detection and protection scheme using integrated features of images, frames and text," *Expert Systems with Applications*, vol. 115, pp. 300-313, 2019.

NPRC Journal of Multidisciplinary Research
Vol. 1, No. 6, Special Issue-I 2024 Pages: 132-141
ISSN: 3059-9148 (Online)
DOI: https://doi.org/10.3126/nprcjmr.v1i6.71754

[4] R. Ronen, M. Radu, C. Feuerstein, E. Yom-Tov, and M. Ahmadi, "Microsoft malware classification challenge," *arXiv preprint arXiv:1802.10135*, 2018.
[5] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in *2009 IEEE symposium on computational intelligence for security and defense applications*, 2009, pp. 1-6.
[6] R. M. Mohammad, F. Thabtah, and L. McCluskey, "Phishing websites dataset," *Available: http://archive.ics.uci.edu/ml/datasets/Phishing+Websites*, 2015.
[7] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. Cambridge, MA: MIT Press, 2016.
[8] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, 2020.
[9] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1153-1176, 2016.
[10] G. Apruzzese, M. Colajanni, L. Ferretti, A. Guido, and M. Marchetti, "On the effectiveness of machine and deep learning for cyber security," in *Proc. 2018 10th International Conference on Cyber Conflict (CyCon)*, 2018, pp. 371-390.
[11] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.
[12] G. Van Houdt, C. Mosquera, and G. Nápoles, "A review on the long short-term memory model," *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5929-5955, 2020.
[13] M. Almiani, A. AbuGhazleh, A. Al-Rahayfeh, S. Atiewi, and A. Razaque, "Deep recurrent neural network for IoT intrusion detection system," *Simulation Modelling Practice and Theory*, vol. 101, p. 102031, 2020.
[14] S. M. Kasongo, "A deep learning technique for intrusion detection system using a Recurrent Neural Networks based framework," *Computer Communications*, vol. 199, pp. 113-125, 2023.
[15] S. Kaur and M. Singh, "Hybrid intrusion detection and signature generation using deep recurrent neural networks," *Neural Computing and Applications*, vol. 32, no. 12, pp. 7859-7877, 2020.
[16] R. H. Chowdhury, N. U. Prince, S. M. Abdullah, and L. A. Mim, "*The role of predictive analytics in cybersecurity: Detecting and preventing threats*," World Journal of Advanced Research and Reviews, vol. 23, no. 2, pp. 1615-1623, 2024.
[17] S. Duary, P. Choudhury, S. Mishra, V. Sharma, D. D. Rao, and A. P. Aderemi, "*Cybersecurity Threats Detection in Intelligent Networks using Predictive Analytics Approaches*," in 2024 4th International Conference on Innovative Practices in Technology and Management (ICIPTM), Feb. 2024, pp. 1-5, IEEE. DOI: 10.1109/ICIPTM59628.2024.10563348
[18] M. Yildirim, "*Artificial intelligence-based solutions for cyber security problems*," in Artificial Intelligence Paradigms for Smart Cyber-Physical Systems, IGI Global, 2021, pp. 68-86.

NPRC Journal of Multidisciplinary Research
Vol. 1, No. 6, Special Issue-I 2024 Pages: 132-141
ISSN: 3059-9148 (Online)
DOI: https://doi.org/10.3126/nprcjmr.v1i6.71754

[19] I. H. Sarker, "*Deep cybersecurity: a comprehensive overview from neural network and deep learning perspective*," SN Computer Science, vol. 2, no. 3, p. 154, 2021.
[20] M. Ahsan, K. E. Nygard, R. Gomes, M. M. Chowdhury, N. Rifat, and J. F. Connolly, "*Cybersecurity threats and their mitigation approaches using Machine Learning—A Review*," Journal of Cybersecurity and Privacy, vol. 2, no. 3, pp. 527-555, 2022.