**International Academic Institute for Science and Technology**

International Academic Journal of Science and Engineering
Vol. 11, No. 2, 2024, pp. 42-46.

ISSN 2454-3896

**International Academic Journal of Science and Engineering**

www.iaiest.com

# An Enhanced Model for Detection of Network Intrusions Using Explainable Artificial Intelligence Through an Enhanced Elman Neural Network

Ankur Dey [a], Deepika Sen [b]

[a] *Digital Twins for Industry Lab, Haldia School of Smart Engineering, India*.
[b] *Digital Twins for Industry Lab, Haldia School of Smart Engineering, India*.

## Abstract

The fourth industrial revolution facilitates innovative business models, including customized manufacturing, continuous surveillance of process conditions and advancement, autonomous choice-making, and online maintenance, among others. Nevertheless, they are more vulnerable to a wide array of cyber-attacks due to constrained resources and their varied characteristics. Such threats result in monetary and social harm to firms, along with the loss of confidential data. This study presents an improved framework for detecting network intrusions using explainable artificial intelligence (DNI-EAI). The Enhanced Elman Neural Network (EENN) model has been employed for DNI, using the Improved Fruitfly Optimization (IFFO) method for optimizing variables. This work used SHAP-EAI technique to enhance the explanation of detection outcomes. The test configuration is established using MATLAB software, involving Honeypot database as inputs. The investigation indicates that the suggested technique attains exceptional performance in DNI, with a detection accuracy of 98.5%.

**Keywords:** Network Intrusions, Elman Neural Network, Fruitfly Optimization, SHAP, Explainable Artificial Intelligence, MATLAB.



DOI: 10.71086/IAJSE/V11I2/IAJSE1149
Received: 24 April 2024/Revised: 18 May 2024/Accepted: 30 June 2024/Published: 30 July 2024

## Introduction and Related Works

The main objective of creating a DNI is to identify many forms of anomalous network activity that a conventional firewall is incapable of detecting. Typical elements of DNI include sensors, a detector, and a user interface. Different types of DNI are used in numerous industrial scenarios (Sharma et al., 2024). There are two categories of safety-based options: preventive tactics and detecting methods. Verification and encryption are used in preventive techniques to safeguard against potential threats (Almuqren et al., 2023; Arreche et al., 2024). When preventive measures are ineffective, detection-based strategies are used. Pattern recognition is a technique of identification. This method utilizes established patterns referred to as anomaly detection for detecting malicious communications. To identify aberrant behavior within the framework using hybrid strategies, these recognition techniques aim to integrate one or more techniques. A dormant DNI detects a breach of security using a DNI detector, logs the incident, and notifies the user interface.

The safety architecture of the commercial manufacturing ecosystem is rapidly evolving. Initially, connecting previously isolated businesses to the Internet suggests that the concept of a "air barrier" no longer safeguards them. Cloud computing and mobile devices are being used in Cyber-Physical Manufacturing Systems (CPMSs) to link diverse parties across the supply chain, hence enlarging the threat space (Sivamohan et al., 2023). Intruders use flaws to gain entry into the network and inflict more severe consequences. Consequently, a reliable DNI is implemented to safeguard the network boundaries, which is essential for facilitating unrestricted interaction without enforcing excessive privilege constraints while allowing diverse components to operate. Various deep learning (DL) approaches have been established over the past years; yet, they are increasingly vulnerable to a wide array of cyber attacks due to restricted resources and their diverse composition (Kima & Euom, 2024). These possibilities can result in economic and social harm to firms, and also the loss of confidential data.

They created the rapid threat identification-based multi-aspect streams of data (MDS-AD) for intelligent DNI in the digital economy. The innovative DNI approach integrates PCA, regional sensitive hash (RSH), and the separation forest methodology. The suggested technique adhered to certain characteristics to get optimal results: (a) multiple aspect information was processed using LSH, (b) the approach forecasted class anomalies based on the findings, and (c) density had been decreased by PCA. The tests used the UNSW-NB15 dataset, yielding the most favorable outcomes. The disadvantage of this approach was the leaking of data during the information minimization procedure.

It elucidated the layered extended learning framework for DNI in the context of Industry 4.0. The approach has been employed to effectively retain the benefits of the learning procedure. This approach used a decode and encode methodology to uncover the latent properties of the network traffic records. The trials were carried out across many practical DNI workloads to achieve maximum effectiveness and outstanding effectiveness in challenging data transmission scenarios.

Authors in (Alohali et al., 2022; Meli et al., 2024) proposed a DL methodology for identifying distributed denial-of-service (DDoS) attacks. Network usage statistics were collected from an operational silicon production facility to forecast the constraints of a current methodology. The database's traffic characteristics were retrieved utilizing the Netmat software. Forty-five bilateral network link characteristics were identified, and several designated databases were created for both learning and testing of the DL system. The study results demonstrated an improvement in reliability and reduced erroneous probability.

## Proposed Method

This work presents an intelligent intrusion detection approach, termed the DNI-EAI technique, for industrial environments. The DNI-EAI approach primarily focuses on the recognition and categorization of intrusions inside the Industry4.0 platform. The DNI-EAI approach encompasses many smaller processes, including data preparation, feature selection, EENN-based detection, and IFFO-based variable adjustment. Figure 1 depicts the comprehensive architecture of the proposed system.

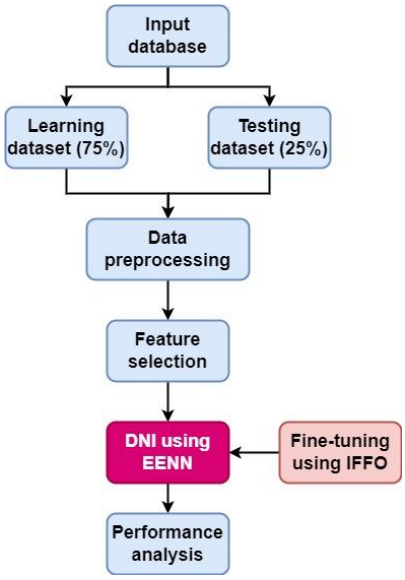



**Fig. 1: Architecture of the Proposed DNI-EAI System**

Information transformation and standardization are the two phases of preprocessing. The input database in excel format is transformed into .csv format during the data transformation. Furthermore, standardization may be executed using the min-max approach, which involves using the lowest and highest values from the specified database. The database is then split into training set (75%) and testing set (25%).

The feature selection technique reduces data complexity by eliminating inconsequential characteristics from the database. It not only eliminates duplication of information but also minimizes the rates of false positives in DNI. Furthermore, it enhances detection rate while decreasing computational burden. Consequently, the feature selection procedure significantly influences the DNI efficiency and adaptation capabilities of the model. This study employed the krill herd optimization (KHO) method for selecting features, which efficiently identifies subsets of attributes by minimizing data dimensionality and enhancing both DNI efficiency and precision.

## DNI-EAI using EENN

The EENN model is used for precise DNI. The ENN was a completely linked adaptive feedback neural network with a regional response function and a regional storage unit. It facilitates the integration of an evolving system and representation of a fixed system, demonstrating the changing characteristics of the model. In contrast to FFNN, Elman functions as an a single-step delay activator that facilitates limited storage and incorporates an additional receptive layer within a three-layer architecture including hidden, input, and output layers. Consequently, the Elman technique is capable of adapting to temporal variations in characteristics. Concurrently, it has robust network reliability and computational capability.

The ENN architecture has one output unit, two input units, and three hidden units. Nonetheless, ENNs include both advantages and disadvantages. Researchers have principally examined and modified the limitations of neural network algorithms in current time engineering contexts, concentrating on two aspects: learning methods and network structure. The modification of neural network topology will significantly exacerbate the computational challenges that hinder the enhancement of neural network performance. Consequently, the momentum methodology enhances the neural network's learning process, indicating that during reverse propagation, a portion of the previous weight modification is included to yield current weight alteration values via velocity components, which are then applied as real weight adjustments. The IFFO methodology has been employed to modify the parameters associated with the EENN method.

The FFO algorithm is fundamentally based on the foraging activities of fruit flies. The IFFO algorithm is characterized by the integration of a stochastic idea. The stochastic signal produced by the predictable system exhibits the characteristic of randomness. The curve is determined by the starting value and the confusion mapping factor. Fitness choice is a critical element of the IFFO algorithm. Response coding may be used to evaluate the merit of the proposed solution. The accuracy score serves as the primary criterion for formulating a measure of fitness.

The EAI methodologies enhance interpretative efficiency by comprehending the effects of harmful data. The SHAP gradient trainer and the extensive SHAP presenter include concepts from the combined gradient method. The kernel-based SHAP method has been chosen for the present investigation. The model-agnostic technique has been employed for the evaluation of SHAP results. The SHAP kernel explainer works for all layouts, although it is slower than alternative class-specific methods. It yields optimal outcomes and delivers precise SHAP values. Every SHAP value delineates the importance that characteristics make to the model's final detection. It enhances the interpretability of the learning approach by illustrating the significance of each attribute. This approach used a collaborative notion for calculating the SHAP scores.

## Results

The test configuration is established using MATLAB software, involving Honeypot database as inputs.

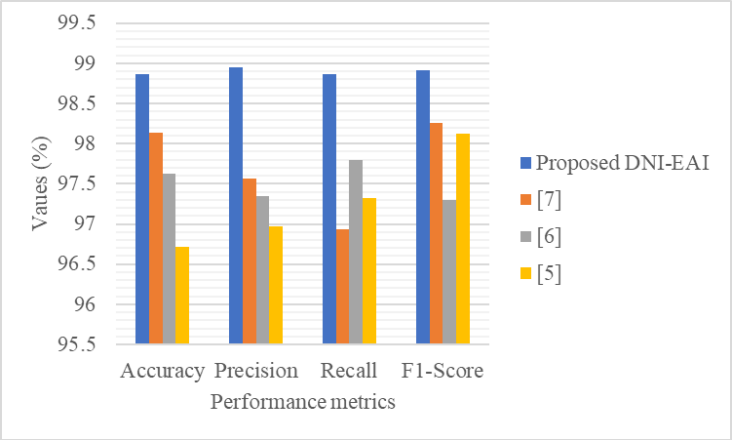



**Fig. 2: Performance Analysis of Proposed DNI-EAI**

Fig. 2 depicts the performance analysis of proposed DNI-EAI. The performance analysis for the proposed DNI-EAI System shows that its accuracy, precision, recall, and F1-score is better than approaches cited in references earlier in the document. The DNI-EAI system has a performance measurement value above 98.5% for all metrics, considerably surpassing other models. Method comes next, but with lower values for recall and some other metrics. Methods seem to have dropped considerably on all metrics. This indicates that with DNI-EAI, better classification effectiveness has been achieved with higher reliability and robustness in performance compared to earlier approaches.

## Conclusion

This paper introduces an enhanced methodology for identifying network intrusions using explainable artificial intelligence (DNI-EAI). The Enhanced Elman Neural Network (EENN) model has been used for DNI, using the IFFO approach for variable optimization. This study used the SHAP-EAI approach to improve the elucidation of detection results. The test setup is constructed with MATLAB software, including the Honeypot database as inputs. The analysis reveals that the proposed approach achieves outstanding performance in DNI, with a detection accuracy of 98.5%.

## References

Sharma, B., Sharma, L., Lal, C., & Roy, S. (2024). Explainable artificial intelligence for intrusion detection in IoT networks: A deep learning based approach. *Expert Systems with Applications, 238*, 121751.

Almuqren, L., Maashi, M. S., Alamgeer, M., Mohsen, H., Hamza, M. A., & Abdelmageed, A. A. (2023). Explainable artificial intelligence enabled intrusion detection technique for secure cyber-physical systems. *Applied Sciences, 13(5),* 3081.

Arreche, O., Guntur, T., & Abdallah, M. (2024). Xai-ids: Toward proposing an explainable artificial intelligence framework for enhancing network intrusion detection systems. *Applied Sciences, 14(10),* 4170.

Sivamohan, S., Sridhar, S. S., & Krishnaveni, S. (2023). TEA-EKHO-IDS: An intrusion detection system for industrial CPS with trustworthy explainable AI and enhanced krill herd optimization. *Peer-to-Peer Networking and Applications, 16(4)*, 1993-2021.

Kima, K. K., & Euom, I. C. (2024). *Research on Implementation of Explainable Anomaly Detection Using Artificial Intelligence on Cyber-Physical Systems.* In KNS Conference Proceedings.

Alohali, M. A., Al-Wesabi, F. N., Hilal, A. M., Goel, S., Gupta, D., & Khanna, A. (2022). Artificial intelligence enabled intrusion detection systems for cognitive cyber-physical systems in industry 4.0 environment. *Cognitive Neurodynamics*, *16*(5), 1045-1057.

Meli, D. (2024, August). Explainable online unsupervised anomaly detection for cyber-physical systems via causal discovery from time series. In *2024 IEEE 20th International Conference on Automation Science and Engineering (CASE)* (pp. 4120-4125). IEEE.