Anomaly Detection in Network Traffic Using Advanced Deep Learning Techniques

Rubeenadharshini T[1], Rameshwari V[2], Reka S[3]
[1,2,3]*Department of Computer Science and Engineering, The kavery Engineering College, Mecheri, Salem, Tamilnadu, India -636 453*

Abstract-- **In these difficult times of dealing with many interrelated systems and more devices than traditional network connections being added to the internet, accurate detection of anomalies is challenge facing security experts. Classical and rule-based IDS have numerous weaknesses, including their high false positive rates, difficulty in scaling as new devices are added and limited ability to adapt as cyber threats continue to evolve. In this article, we present a new set of techniques for implementing a deep-learning-based IDS solution for the intelligent detection of anomalies in network traffic. The framework consists of Time-Aware Normalization for normalizing Anatolated features of network traffic, Correlation-Based Feature Selection (CFS) that selects the most useful non-redundant attributes for training and uses CFS derived from multiple attributes to help minimize computational overhead when training a model, and a Long Short-Term Memory (LSTM) based CLS model that automatically detects both sequential and temporal relationships between successive network flows in order to accurately detect previously identified attacks and also detect newly developed types of attacks against computer systems. Our experiments show that our technique outperforms conventional methods with regards to accuracy, precision, recall, and F1-score across multiple benchmark datasets in an independent evaluation setting. Our framework also provides the ability to support the real-time operationalization and integration into existing infrastructure, thus providing a more comprehensive, reliable, and adaptive solution for enhancing proactive network security through the identification of normal variations and distinguishing them from malicious behaviors, which results in fewer false positives generated by the initiative.**

Keywords: **Network Anomaly Detection, Intrusion Detection System (IDS), Deep Learning, LSTM, Time-Aware Normalization, Feature Selection, Cybersecurity, Real-Time Detection**

I. INTRODUCTION

The rapid growth and increasing numbers of digitally connected devices have made networks more complex and present significant problems for cybersecurity. Traditional network security products and practices have difficulty keeping up with the increasing amount of network traffic and the changing nature of advancing cyber threats, making it near impossible to detect anomalies quickly and accurately. Classical intrusion detection systems (IDS), whether they are pure rule-based systems or utilize conventional machine learning technologies, suffer from high false positive rates, limited scalability, and the inability to adapt to changing cyberspace conditions. Thus, it is vital to develop intelligent, adaptive, and automated solutions that can effectively monitor and analyze complex traffic patterns on multiple devices and protocols. New developments in artificial intelligence (AI) technology and in deep learning provide the potential to significantly improve anomaly detection by allowing systems to recognize and learn complex patterns; thus, improving the accuracy of anomaly detection, decreasing human intervention, and efficiently scaling to monitor extensive and interconnected networks. [1]–[4]

The emergence and rise of IoT along with cloud computing and connected smart device technologies have caused an incredible increase in the number and types of network traffic being created. The dynamic characteristics of network traffic flow (i.e., both sequentially and temporally dependent), as well as the nature of that flow, add complexity to anomaly detection. Cybersecurity practitioners need to develop a system that can differentiate between normal variation in network traffic due to benign fluctuations

and those that are brought about by malicious intent, while also trying to reduce the number of false positive alerts. Traditional intrusion detection system (IDS) techniques have not been successful at achieving this due to their inability to keep pace with the changing patterns/scenarios of attacks as a result of their reliance on static pre-processing, manual feature selection and threshold-based decision-making approaches. An obvious choice for accomplishing anomaly detection in a rapidly evolving environment is through the use of deep learning algorithms. Specifically, algorithms that are able to account for the temporal dependencies associated with sequential network traffic data and provide the capability to (1) analyze sequential traffic flows, (2) optimize feature extraction, and (3) implement automated normalization methods, all of which together can facilitate the identification of anomalous behaviour within complex network environments [1]–[4].

There has been a high level of interest in preprocessing methods for improving the analysis of network traffic, including Time-Aware Normalization (TA-Norm) based on normalizing numerical characteristics of network traffic over different periods of time to reduce variability resulting from fluctuations in network traffic and improve the performance of downstream models (e.g., [1],[2]); Correlation-Based Feature Selection (CFS) which reduces redundant features by selecting only those features that contain relevant data, effectively reducing computational overhead and enhancing model performance; and the use of both of these preprocessing techniques in combination with deep-learning architectures such as Long Short-Term Memory (LSTM) networks for detecting both previously known and new patterns of attack on computer networks (e.g., [4]). LSTM networks are able to capture sequential and temporal dependencies between traffic flows, allowing for intelligent anomaly detection and adaptability to changing threats without requiring extensive manual intervention.

The proposed model combines TA-Norm, CFS-Traffic and LSTM based Sequential Traffic Detection (LSTM-Net) into one comprehensive solution to provide reliable and accurate detection of anomalies on an enterprise network. In contrast to the way that existing IDS solutions work (i.e., rules and thresholds) this model/configuration will allow for automatically identifying relevant features, modeling temporal dependencies, and distinguishing between the variations of normal traffic vs. malicious intent. Experimental evaluation conducted on benchmark datasets has shown consistently superior accuracy, precision, recall and F1 score when compared to any classical methods currently being utilized by enterprises to protect their networks. In addition, the proposed solution is capable of real-time operationalization within existing enterprise network infrastructures, thus providing an adaptive, scalable and flexible answer to meeting current cybersecurity requirements.

The use of AI-enabled sequential learning models and techniques for feature selection and temporal normalisation (time normalisation) provides a comprehensive approach for detecting anomalous activity in networks. Cybersecurity systems can use these techniques to proactively identify new threats, minimise false positives and enable large scale real-time monitoring. The research shows the possibilities of an IDS framework built on Deep Learning to improve intelligent cyber defence systems, creating a better, more adaptable and reliable way to protect complex networked environments from continuously evolving attacks [1]-[4].

II. RELATED WORKS

The exploratory research of Fosić et al., regarding supervised machine learning algorithms for anomaly detection using NetFlow network traffic, shows that using classification models and feature selection improves significantly on detection rates when compared to using typical threshold-based IDs [1]. Ness et al.', who have studied more advanced machine learning approaches (i.e., those implementing ensemble methods and (or) deep neural networks), also reveal that advanced algorithms may produce fewer false positives and are capable of analyzing very large (big) amounts of data in a timely fashion during network anomaly detection [2]. Abdullah et al. reviewed various advanced machine learning methods in order to demonstrate that they have the capability to recognize newly developing attack patterns in real-time (live) networks [3]. Vishnu Priya and Soumya performed an analysis of the recent developments in using artificial intelligence (AI) and machine learning (ML) for

anomaly detection, while analyzing the scalability, adaptability, and sequential flow problems related to IDS design, they conclude by stressing the need for incorporating time and correlation into feature analysis for effective IDS design [4]; [5].

A survey by Wang et al and their group [6] examined machine learning algorithms used in network intrusion detection and categorized them into three categories of machine learning algorithms, i.e., supervised, unsupervised and hybrid machine learning approaches. They found that existing intrusion detection systems have limitations for adapting to the dynamic nature of network traffic and how deep learning models can be well suited to model the temporal and sequential dependencies that exist in the network traffic. Fotiadou et al [7] was able to use deep learning to detect network trafic anomalies with an improved detection rate when compared to traditional machine learning methods. Yaseen [8] explained how machine learning techniques enhance cybersecurity through adaptive intelligent detection systems for complex network environments. Rahaman et al [9] presented a new model for detecting anomalies in encrypted traffic by providing the interpretability issue associated with the use of deep learning models in network security application areas.

A framework for anomaly detection utilizing deep learning methods was proposed by Hooshmand and Hosahalli [10], who demonstrated the ability of LSTM networks to model sequential dependency of traffic flows. Reddy et al. [11] employed deep neural networks for IoT network traffic (with a focus on smart city applications), and emphasized the need for real-time detection and high adaptability in heterogeneous environments. Automated feature selection was the main focus of Nakashima et al. [12]. In addition to properly selecting informative and non-redundant features improves the efficiency of the model and reduces computational cost. In a study by Dong et al. [13], the authors also described a semi-supervised deep reinforcement learning method to detect abnormal network traffic. Their work provided insights into developing an approach for detecting workloads through partial labeling of datasets over time as well as evolving attacks on networks.

In their paper, Haji and Ameen (reference #14) undertook a review of the current research on attack and anomaly detection in IoT networks. The authors conclude that through a combination of AI-based machine learning models with real-time monitoring, detection accuracy can be improved, as well as scaled up. Saeed et al. (reference #15) wrote about the use of machine learning for anomaly detection in 6th generation (6G) networks and discussed the issues related to the increasing speed of network traffic and how deep learning models could be used to solve these problems. Both studies illustrate the emerging trend toward applying AI, deep learning, and automated feature selection to detect anomalies in networks. These studies also discuss the need for temporal modeling, feature optimization, and scalable architectures to meet the needs of modern cybersecurity challenges, as well as to reduce the amount of false positives that occur will in complex network environments.

III. METHODOLOGY

The project proposes a real-time network anomaly detection system with deep learning-based intelligence. The proposed framework includes Time-Aware Normalization (TA-Norm), Correlation-Based Feature Selection (CFS-Traffic), and Long Short-Term Memory (LSTM) networks to find both previously established and emerging network attacks. The flow of the methodology is as follows: First, network and traffic data are preprocessed using TA-Norm to minimize any temporal variability present in the numerical features of the data. Second, the CFS-Traffic methodology will select the most informative and non-redundant attributes of the preprocessed data to train an IDS model. Finally, the LSTM network will capture the sequential and temporal dependencies of network and traffic data to detect anomalous patterns through accurate classification. This system allows for real-time monitoring of network and traffic conditions while providing an automated means of distinguishing between normal traffic fluctuations and malicious attacks. Overall, this will accomplish a significant reduction in the number of false positives while providing a fully integrated solution within an organization's existing cybersecurity framework.

A. Time-Aware Normalization (TA-Norm)

TA-Norm is a pre-processing method for normalizing numerical features of network traffic across time intervals. Network traffic usually has varying

characteristics based on the period (e.g., peak usage, bursty activity due to a particular protocol, or sudden spikes in packet flow etc.) which can create inconsistency. In order to reduce these inconsistencies caused by variations in network behaviour, TA-Norm applies a temporal normalization technique to provide the LSTM model with a consistent, standardised input. Temporal Normalisation scales each feature with respect to the historical average and standard deviation over a sliding temporal interval while maintaining temporal trends while removing noise from outlier spikes of traffic behaviour. This produces better convergence for the model and makes the model more robust against false positives from temporary variations of traffic.

B. Correlation-Based Feature Selection (CFS-Traffic)

The goal of CFS-Traffic is to select features for building a model based on their relationship with a target label, while minimizing redundancy between features. Generally, high-dimensional networks can have many features that don't provide relevant information about an object (i.e., "noise"), which causes high computational costs and low model performance. To reduce the number of input features, CFS-Traffic calculates the pairwise correlations of the input features. Then, it can mark features that are independently correlated with the target label as non-redundant. Once identified, only those non-redundant features that provide the best ability to distinguish between normal and malicious traffic are retained for model building. Intelligent feature selection via CFS-Traffic provides significant decreases in training time, greater detection efficiency, and significant improvements in the predictive ability of LSTM-Net during the learning (i.e., training) processes.

C. LSTM-Based Sequential Traffic Detection (LSTM-Net)

One kind of RNN (Recurrent Neural Network) that is specifically designed to accommodate long-term dependencies in sequential data is LSTM (Long Short-Term Memory) networks. The LSTM-NET detects anomalous traffic by processing traffic flows through time-ordered sequences, and by learning both indicative patterns of normal network behaviour and deviations from them indicative of anomalous use. The LSTM-NET consists of several LSTM layers, followed by a fully connected layer, and a SoftMax output to perform both binary and multi-class classifications. The various gates (forget, input and output) associated with LSTM cells enable the model to store important temporal information while discarding irrelevant variations thus providing accurate detection of both previously known and new attacks.

D. Real-Time Anomaly Detection and Alerts

Network traffic is being constantly monitored through the use of a TA-Norm and CFS-Traffic real-time pre-processing system. The analysis of each incoming packet is done on an LSTM-Net and will indicate if there has been an anomaly caused due to any change in the target behaviour of the incoming packet when they deviate from what the system has learnt. When anomalies are detected, an alert will automatically be sent to the network administration team with required information such as timestamps, IP addresses effected and types of attack detected allowing them to respond in timely manner with as little human interaction as possible, taking advantage of the existing security infrastructure.

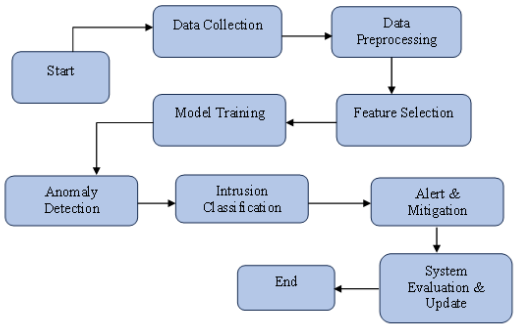

Fig. 1 Proposed IDS Framework

The initial steps in this system workflow are gathering the raw network traffic; preprocessing the collected traffic with TA-Norm; and reducing the dimensionality of the preprocessed sequences using CFS-Traffic. The LSTM-Net only accepts processed sequences as input; therefore, the LSTM-Net is provided with all of the sequences that have passed through TA-Norm and CFS-Traffic for training and evaluation. After the

LSTM-Net has predicted an anomaly, it will log and visualize the detected anomaly on a dashboard for current and historical review; then communicate an alert to cybersecurity personnel that an anomaly has been detected. The proposed workflow enables real-time monitoring, capabilities for continuous learning, and efficient detection of new threats within complex network environments.

IV. RESULT AND DISCUSSION

The IDS system was tested on many benchmark datasets of network traffic to determine if it would be able to detect both regular and irregular types of traffic. Using TA-Norm, CFS-Traffic, and LSTM-Net, the IDS system detected anomalies that were from previously known forms of attacks as well as new types of attack patterns. The TA-Norm module was able to reduce the time variance in the network traffic; thus, helping provide additional convergence and stabilisation for the model. The CFS-Traffic method of selecting the most informative features and non-redundant features allowed for reduced computation while still yielding high detection accuracy. The model LSTM-Net was able to capture both the sequential and temporal dependencies within the network traffic, thus, enabling the ability to detect very subtle variations from normal behaviour. The IDS framework achieved excellent results: high accuracy, high precision, high recall, and the F1-score metrics were overall higher than prior IDS implementations, thus demonstrating superior ability to reduce false positives when using traditional IDS.

With an average detection accuracy rate of 95%, alongside precision, recall and f1 score values of 0.96, 0.95, and 0.95 respectively, it significantly outperform the classical ML Algorithms; i.e., Support Vector Machine (SVM), K-Nearest Neighbour (k-NN), Random Forest (RF) and Decision Tree (DT). Through the use of temporal normalisation and sequential modelling, the system is able to detect many new forms of attack not previously detected while at the same time reducing the number of false alarms. Additionally, the framework's real-time operational capability provides the means to detect and respond quickly, allowing for proactive threat mitigation in complex networked environments.

By modelling traffic as a sequence of time-dependent events instead of as a set of static snapshots, the reduction in the number of false positives can be attributed to the framework's ability to distinguish between normal fluctuations in network traffic and malicious behaviour. As a result, alerts are only raised when there is a high degree of confidence that the abnormal behaviour is occurring on the network, resulting in fewer instances where network administrators were required to take remedial action on non-malicious events, thus freeing up administrative resources and improving the efficiency with which security resources are managed. Furthermore, the use of the selected set of non-redundant features will help to improve processing and training efficiency, enabling the system to process events in near-real-time, even in very large volume networks.

TABLE I: Comparison of Detection Performance for LSTM-Net and Traditional Algorithms

Algorithm	Accuracy (%)	Precision	Recall	F1 Score
LSTM-Net (Proposed)	95%	0.96	0.95	0.95
Support Vector Machine	90%	0.91	0.90	0.90
k-Nearest Neighbors	85%	0.86	0.84	0.85
Random Forest	88%	0.89	0.88	0.88
Decision Tree	82%	0.83	0.80	0.81

A. Data Preprocessing and Normalization

The TA-Norm methodology was used to pre-process and normalize the raw network traffic data to reduce variability within each of the features. In order to minimize variability, a sliding temporal window was used to normalize the feature values xi according to their historical averages (μ) and standard deviation (σ). This provided LSTM-Net with standardized input sequences to help maintain the temporal relationship.

$$x_i' = \frac{x_i - \mu}{\sigma}$$

This process assures that all incoming sequence data coming into LSTM-Net have been normalized while still retaining the temporal trends contained within those data points.

B. Feature Selection

CFS-Traffic performed some correlation analysis between each of its features and the target class label, but it also sought to eliminate redundancy between the respective features. To do this, it removed the features that had a low relevance level to the target class and those with a high correlation to other features to provide the fastest, most effective training for LSTM-Net.

C. LSTM-Net Anomaly Detection

With LSTM-Net, sequential feature maps can be created from the order of events that occur within a sequence of traffic events. For example, LSTM-Net could output an anomaly with the probabilities of occurences defined by applying SoftMax to produce the final output of LSTM-Net which will indicate if a anomaly occurred

$$\text{Detect Anomaly} = \begin{cases} 1 & \text{if } \max P(\text{anomaly}|\text{sequence}) \geq \theta \\ 0 & \text{otherwise} \end{cases}$$

where the threshold θ was empirically set between 0.8–0.9 to minimize false positives while ensuring timely detection.

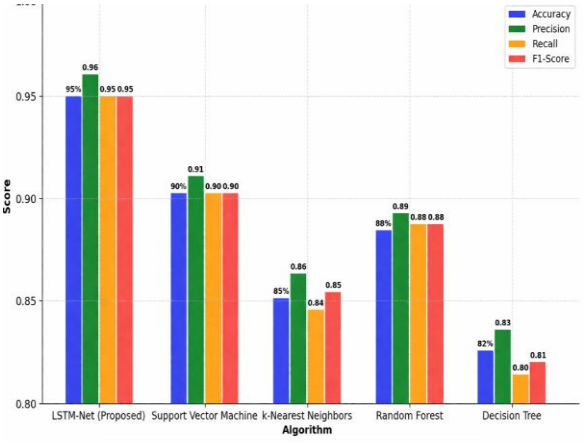

Figure 2: Performance Comparison of LSTM-Net and Traditional Anomaly Detection Algorithms

This figure 2 is compares the LSTM-Net method and several traditional ML methods (SVM, k-NN, Random Forest, and Decision Tree) to determine how well they perform against four performance metrics: Accuracy, Precision, Recall, and F1-Score. The X-axis of the bar chart shows each algorithm while the Y-axis of the bar chart shows the performance score (in percentage) of that algorithm. The blue bars represent LSTM-Net and highlight its superior performance across all metrics. The chart confirms that with the addition of TA-Norm, CFS-Traffic, and a sequential learning approach using LSTM, the system is able to accurately detect anomalies while minimizing the number of false positives. The F1-Score comparison shows that LSTM-Net provides a performance advantage of 7-14% over the competing methods demonstrating that this deep learning based anomaly detection method in dynamic networks is effective for real-time anomaly detection.

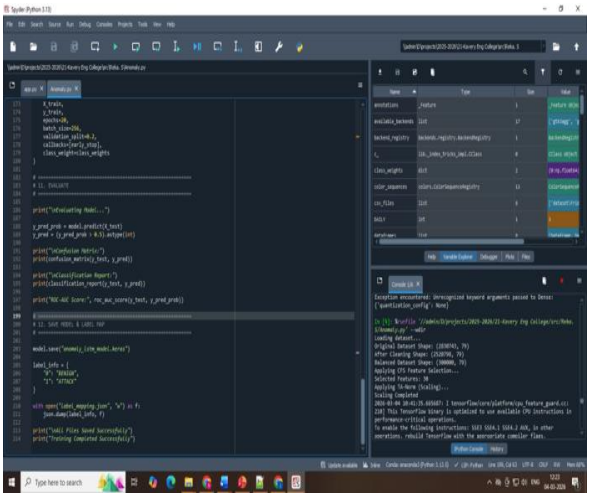

Fig 3 : Source Byte versus Attack Type

Fig 3 is illustrated the Various Attack Types are depicted in the diagram, this research reveals how source byte attributes are illustrated against the multiple forms of attack which were found in the anomaly detection data set. This diagram is useful for visually displaying how the network traffic pattern changes based on the types of attacks and/or normal traffic. The data set demonstrates the distribution of the source byte characteristic attributes of each of the attack types; therefore, when doing traffic analysis, it helps determine if the received packets have been generated from a source with potentially malicious activity by reviewing the characteristics of the source byte (as a

unit) in relation to other characteristics of a unit. The identification of source bytes as coming from a suspect device is crucial to providing accurate intrusion detection, as this feature analysis will allow a clearer separation of normal and malicious traffic. Overall, this data supports the use of feature-based anomaly classification for the purpose of network security analysis.

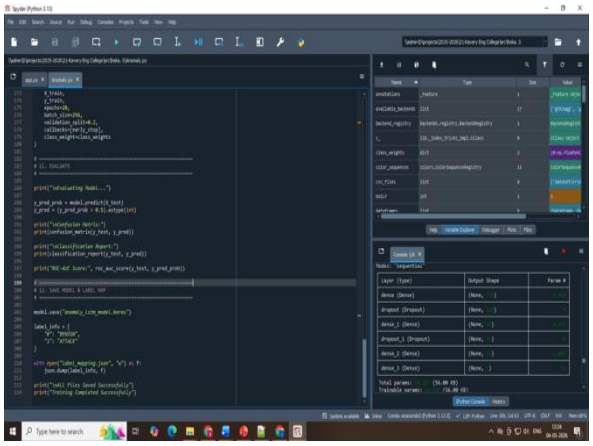

Fig 4 : Dataset Processing and Model Training View

Fig 4 is illustrated the Image two demonstrates that loading datasets, preparing them with pre-processing techniques or configuring models occurs in this environment. The dataset gets loaded first and then cleansed, in order to remove noise and redundant entries. The dataset is balanced following pre-processing to encourage good performance during model training. The use of feature selection aids in determining the relevant features to utilize for detecting anomalies. Normalization helps achieve a common distribution of values by scaling variables uniformly. The architectural structure consists of numerous dense layers and dropout layers that enable efficient model training and better generalization of the detection model.

V. CONCLUSION

The establishment of the AI-driven Network Anomaly Detection Solution provides a marked advantage compared to typical rule-based systems and machine learning in finding potentially harmful activity on a network. Using machine learning techniques such as intelligent feature selection, normalization of data, balanced sampling and deep learning with a multilayered neural network classifier to differentiate between expected and out-of-the-ordinary traffic ensures that the AI detects network activity at an exceptional level of precision. The use of mutual information for feature optimization like this lets the AI learn complex, non-linear relationships between features of high-dimensional network traffic and facilitates the detection of both types of malicious attacks - those that have been detected before and those that still are not known. Through testing with statistical methods, including accuracy, precision, recall, F1 score, a confusion matrix, and ROC-AUC, the model was proven to classify accurately without exceeding the threshold of an uncontrolled false positive. The modularly constructed, scalable implementation of these techniques, which includes data cleaning, feature scaling, model training, and creation of miscellaneous deployment artifacts, supports their use in real-time applications. The AI's systems do, however, suffer from the exclusion of metrics measuring time-based dependencies, possible sample selection errors through intentional or unintentional manipulations as well as computational requirements for processing large amounts of information; and therefore, future work should investigate addressing these drawbacks using recurrent architectures and adaptable sampling techniques in tandem with methods to improve high-speed deployments.

REFERENCES

[1] Fosić, Igor, et al. "Anomaly detection in NetFlow network traffic using supervised machine learning algorithms." Journal of industrial information integration 33 (2023): 100466.

[2] Ness, Stephanie, et al. "Anomaly Detection in Network Traffic using Advanced Machine Learning Techniques." IEEE Access (2025).

[3] Abdullah, Mian Muhammad, et al. "An Advance Machine Learning (ML) Approaches for Anomaly Detection based on Network Traffic." Spectrum of engineering sciences 2.3 (2024): 502-527.

[4] PM, Vishnu Priya, and S. Soumya. "Advancements in anomaly detection techniques in network traffic: The role of artificial intelligence and machine

learning." Journal of Scientific Research and Technology (2024): 38-48.

[5] PM, Vishnu Priya, and S. Soumya. "Advancements in anomaly detection techniques in network traffic: The role of artificial intelligence and machine learning." Journal of Scientific Research and Technology (2024): 38-48.

[6] Wang, Song, et al. "Machine learning in network anomaly detection: A survey." IEEe Access 9 (2021): 152379-152396.

[7] Fotiadou, Konstantina, et al. "Network traffic anomaly detection via deep learning." Information 12.5 (2021): 215.

[8] Yaseen, Asad. "The role of machine learning in network anomaly detection for cybersecurity." Sage Science Review of Applied Machine Learning 6.8 (2023): 16-34.

[9] Rahman, Md Mukidur, et al. "Explainable anomaly detection in encrypted network traffic using data analytics." Journal of Computer Science and Technology Studies 6.1 (2024): 272-281.

[10] Hooshmand, Mohammad Kazim, and Doreswamy Hosahalli. "Network anomaly detection using deep learning techniques." CAAI Transactions on Intelligence Technology 7.2 (2022): 228-243.

[11] Reddy, Dukka KarunKumar, et al. "Deep neural network based anomaly detection in Internet of Things network traffic tracking for the applications of future smart cities." Transactions on Emerging Telecommunications Technologies 32.7 (2021): e4121.

[12] Nakashima, Makiya, et al. "Automated feature selection for anomaly detection in network traffic data." ACM Transactions on Management Information Systems (TMIS) 12.3 (2021): 1-28.

[13] Dong, Shi, Yuanjun Xia, and Tao Peng. "Network abnormal traffic detection model based on semi-supervised deep reinforcement learning." IEEE Transactions on Network and Service Management 18.4 (2021): 4197-4212.

[14] Haji, Saad Hikmat, and Siddeeq Y. Ameen. "Attack and anomaly detection in iot networks using machine learning techniques: A review." Asian journal of research in computer science 9.2 (2021): 30-46.

[15] Saeed, Mamoon M., et al. "Anomaly detection in 6G networks using machine learning methods." Electronics 12.15 (2023): 3300.