JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES
www.journalimcms.org

ISSN (Online) : 2454 -7190 Vol.-19, No.-7, July (2024) pp 17-27 ISSN (Print) 0973-8975

DESIGN AND ANALYSIS OF INTRUSION DETECTION SYSTEM USING MACHINE LEARNING IN SMART HEALTHCARE SYSTEM

K. S. Yamuna[1], M. Sugumaran[2] , A. Arthi [3], R. Premkumar[4]

[1,2] Electrical and Electronics Engineering, Sona College of Technology, Salem, Tamilnadu, India

[3] Department of Artificial Intelligence and Data Science, Rajalakshmi Institute of Technology, Chennai, Tamilnadu, India

[4] Department of Electrical and Electronics Engineering, Sri Eshwar College of Engineering, Coimbatore, Tamilnadu, India

Corresponding Author: **K. S. Yamuna**

Email: [1]yamunaks@sonatech.ac.in, yamuunaks@gmail.com

https://doi.org/10.26782/jmcms.2024.07.00002

(Received: May 17, 2024; Revised: June 22, 2024; Accepted: July 04, 2024)

Abstract

The integration of the Internet of Things (IoT) in medical applications into healthcare applications has enabled the remote monitoring of patients' information, facilitating timely diagnostics as required. The technology of the Internet of Medical Things (IoMT) empowers doctors to treat patients through real-time monitoring and remote diagnostics. Nevertheless, implementing high-security features that ensure the accuracy and confidentiality of patients' data poses a substantial challenge. IoMT devices have limited processing power and memory, making it impossible to build security technology on them. Methodology: So the proposed work formulates a machine learning-based topology to construct an efficient and precise intrusion detection system using network traffic and patient data. Findings: In this topology, modified Whale optimization topology has been implemented for feature selection, and the intrusion is detected using two ML algorithms namely, Random Forest and SVM. Hence, the proposed method surpasses the current state-of-the-art, achieving an accuracy rate of 99.82%.

Keywords: Intrusion Detection System (IDS), Network Attacks, SVM, Random Forest (RF), Modified Whale Optimization Algorithm (MWOA).

I. Introduction

IoMT holds significant promise in revolutionizing healthcare by enhancing health monitoring, and management and ultimately contributes to improved quality of life. The World Health Organization (WHO) has raised concerns about a looming shortage of healthcare professionals by the year. This impending shortage underscores the necessity of the development of cost-effective healthcare architecture

and solutions aimed at preventing diseases, reducing healthcare expenditures, and ensuring a high quality of life for individuals. The IoMT refers to the IoT incorporation with medical contrivances. Nearly 30% of the market for IoT devices is occupied by IoMT devices. Despite the significant research and industry focus on the IoMT and its beneficial advantages, there remains a critical gap in establishing robust security measures for IoMT systems[I, II]. The urgency to rapidly develop IoMT products due to fierce competition among vendors has led to the proliferation of nonstandard devices that utilize varied communication protocols and data transfer topology. These complexities give rise to serious security, privacy, and authentication issues.

In IoMT, several approaches and processes may be utilized to identify and mitigate attacks. Detection and prevention techniques include database surveillance, susceptibility control, hazard information, final equipment surveillance, penetration testing, and preventative systems. In IoMT, network attacks and security vulnerabilities are frequently detected using the intrusion detection system (IDS). Thus, to detect network breaches in IoT systems, the IDS uses network flow abnormalities, signature-based procedures, or protection regulations. Due to the suspect's ongoing use of complex hacking methods and attack strategies, many conventional security detection approaches are ineffective. In recent times, researchers have shifted their focus toward IDS that leverage Machine Learning (ML) and Deep Learning (DL) algorithms [III-VI]. This shift is driven by the unique ability of these algorithms to effectively detect zero-day attacks, mitigate security threats, and adaption to dynamic changes in the IoMT environments. Moreover, IDS based on AI algorithms exhibit promising outcomes in addressing challenges inherent to IoMT, such as system heterogeneity, latency, and scalability issues.

Hence, this work formulated a novel topology using an ML algorithm for IDS in IoMT healthcare devices.

Ultimately, the main accomplishments provided by this work are as follows.

- To effectively detect IoMT attacks, develop an efficient ML-based framework using the combined datasets of patient biometrics and network traffic.
- Achieve a higher accuracy of 99% when correlated with the recent efforts that have been reported on the same datasets.

II. Related Works

Recently, IDS using ML have recently been proposed for IoMT networks. A comprehensive analysis of the continuous training approach was carried out and put into practice by leveraging datasets obtained from the Modbus communication protocol and HTTP networks [XI]. This topology utilized the CIC-IDS2017 dataset to assess the effectiveness of ML models. In this model, the decision tree classifier achieved an impressive accuracy (96.44%) for classifying attacks. Conversely, the RF achieved 94.45% accuracy when applied to the Modbus datasets.

An advanced ML technique, a Deep Recurrent Neural Network (DRNN) along with ML models such as RF, KNN, Decision Tree, and Ridge Classifier has been formulated to find IoMT attacks. To optimize the feature selection process, the PSO algorithm is employed. From the findings, it can be concluded that this proposed

model achieved an outstanding accuracy (about 99.76%) in classifying cyber threats in an IoMT environment [XIV].

A novel Swarm-Neural Network (Swarm-NN) approach has been formulated to discern potential attackers within the IoMT framework. To effectiveness of the proposed Swarm-NN strategy, is verified using real-time data from a secure dataset, specifically the ToN-IoT dataset. Its performance was then compared against conventional classification models using a variety of performance metrics. Thus, it achieved an outstanding accuracy of 99.5%. This noteworthy accuracy underscores the efficacy of the Swarm-NN strategy in identifying and mitigating attacks in the context of data transmission within the IoMT framework [X].

A novel framework for cyberattack detection in IoMT networks, leveraging Ensemble Learning (EL) and a fog-cloud architecture has been proposed. This EL integrates Naive Bayes, Decision Tree, and RF as primary learners in the first level. In the subsequent level, XGBoost is utilized to distinguish between attack and normal cases effectively. In this, a deployment architecture that incorporates Software as a Service (SaaS) within the fog computing domain, complemented by Infrastructure as a Service (IaaS) on the expansive cloud infrastructure. This approach has undergone rigorous validation using the ToN-IoT dataset, a meticulously curated collection derived from a sprawling and diverse IoT network. The experimental results show the remarkable efficacy of this framework, attaining an outstanding detection rate (99.98%) and an accuracy level (96.35%) [VII].

A real-time Enhanced Healthcare Monitoring System (EHMS) testbed was designed to monitor biometric readings of patients and capture essential network flow metrics. Particularly for man-in-the-middle attack (MITM), this system has generated a dataset comprising over 16,000 records encompassing both attacked and normal healthcare data. To bolster the system's resilience against these attacks, various ML methods have been adopted. The results are promising and depict a notable performance enhancement over IDS [V].

A novel intrusion detection system leveraging a DL-based approach, Deep Belief Network (DBN) algorithm model for IDS in IoMT network. The performance analysis of this IDS focuses on attacks and abnormality detection, utilizing the dataset CICIDS 2017. These results underscore the efficacy of the proposed DBN algorithm model in enhancing intrusion detection across diverse attack categories [IX].

Moreover, upon a comprehensive review of the existing literature, it becomes evident that DL models predominantly rely on network connectivity patterns and statistical information to categorize or detect cyberattacks [VIII]. Hence, this work evaluates and compares the performance of two ML methods for the detection of attacks in the IoMT network.

III. Methodology

The IoMT IDS architecture contains various components, including sensor devices, an IDS, an IoT gateway, and security operators responsible for monitoring and responding to potential attacks. The sensor devices encompass a range of healthcare monitoring equipment such as temperature sensors, heart and pulse rate detectors, ECG devices, BP and respiration rate monitoring devices. Communication within the IoMT ecosystem can be facilitated through IoT network protocols such as

MQTT (Message Queuing Telemetry Transport). Figure 1 depicts the architecture of IoMT. The IoT gateways gather sensor data via communication channels and subsequently transmit this data to remote locations.

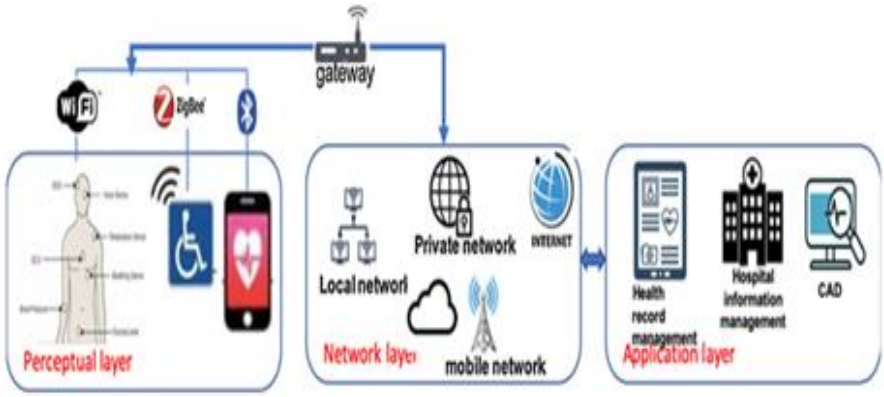

Fig 1. IoMT architecture

Figure 2 represents the design of the proposed topology which is utilized to predict the attacks over IoMT.

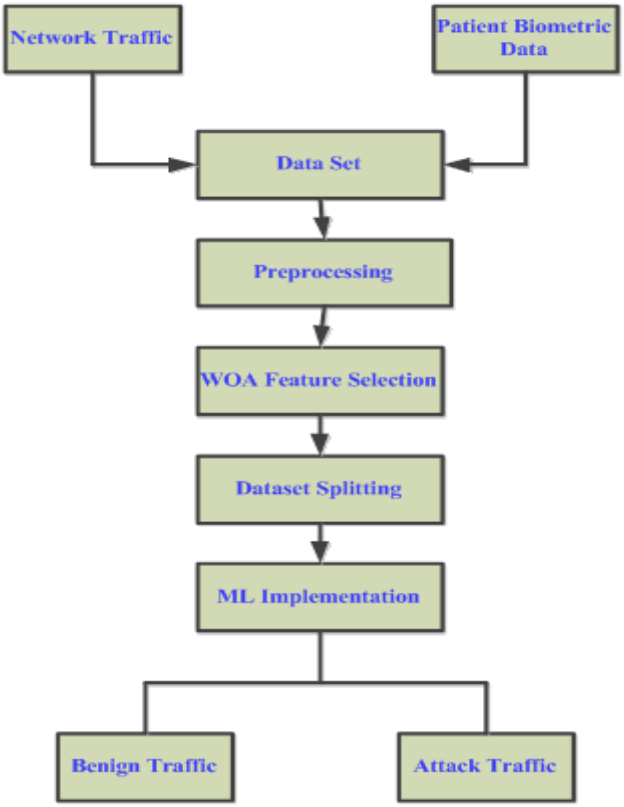

Fig 2. Proposed architecture –Illustration

A. Dataset

The WUSTL-EHMS-2020 dataset was meticulously curated from a live EHMS testbed [5]. This sophisticated testbed is capable of concurrently gathering crucial data streams, encompassing both comprehensive network flow metrics and detailed biometrics of patients. It encompasses various cyberattack types, including data injection, spoofing, and MITM attacks. Statistical details regarding the WUSTL-EHMS-2020 dataset are presented in Table 1. The dataset encompasses a comprehensive set of 44 features, categorized into 35 related to network flow data, 8 derived from patient biometric information, and one designated as the label feature for further analysis.

The dataset is categorized into two classes: "attack" and "normal traffic." Specifically, attack traffic is denoted by the label "0," while normal traffic is indicated by the label "1." As a result, the dataset is composed of 14,272 instances of normal traffic and 2,046 instances of attack traffic. To create a representative subset for analysis, randomly selected 1,000 instances from the attack traffic category is utilized for further investigation.

Table 1: Dataset Splitting

Dataset	Raw Data		Selected Data	
	Normal	Attack	Normal	Attack
WUSTL-EHMS-2020	14,272	2046	14000	1000

Thus, the obtained data set is divided into 65% training and 35% test data for experimental evaluation. This comprehensive setup was utilized to simulate and study Man-in-the-Middle attacks.

B. Preprocessing

Before being employed for IDS, the data underwent a preprocessing phase. This process involved the application of a Standard Scalar methodology to normalize the input data. Utilizing this normalization technique, resulted in transforming the feature data to conform to a distribution with a mean of zero and a variance of one.

C. Feature Selection

Feature selection is a crucial step aimed at enhancing both the accuracy and the speed of prediction. Hence, from the preprocessed data, numerous features have been derived using the Whale Optimization Technique (WOA). Let the features of the dataset be represented as f1,f2,f3 …fn. Feature selection is typically represented using a binary format where a value of "1" denotes the selection of a feature, while "0" indicates that the feature is not chosen. This binary representation helps streamline the feature selection process. This work adopted a modified WOA for feature selection. Figure 3 illustrates the flowchart representation of the modified WOA-based feature selection process.

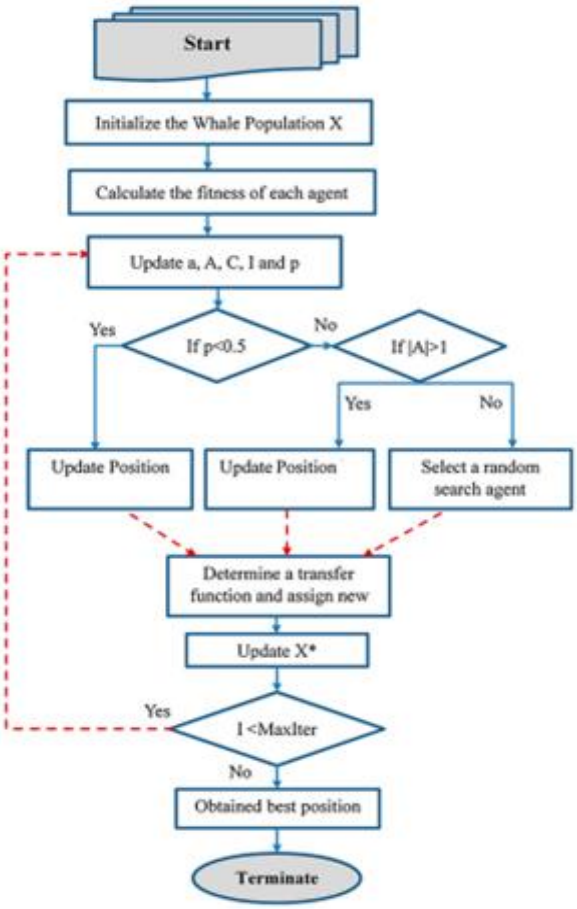

Fig 3. Feature selection Process

The devised fitness function in the proposed methodology aims to strike a balance between two key objectives namely:

- Reducing the quantity of chosen features in every solution.
- Maximizing the classification accuracy achieved through the utilization of these chosen features.

Thus, fitness can be calculated as follows.

Fitness= 〖αγ〗 _R (D)+β |R|/|C|

Where

γ_R (D) - Classification Error Rate

|R| - Cardinality of the subset

|C| - Features in the dataset

α - Classification quality and subset length.

α ϵ[1, 0] and β =(1 - α).

After extracting the required features, the attacks can be classified using ML-based classifiers [XII, XIII]. The subsequent ML algorithms are devised in this work for intrusion detection.

D. Classifiers

Random Forest (Rf)

RF stands as a widely embraced ML technique, exhibiting exceptional flexibility and robustness to adeptly manage both regression classification and classification tasks. This EL method amalgamates numerous decision trees to formulate predictions and conduct data analysis.

Support Vector Machine (SVM)

SVM stands out as a potent classification algorithm extensively employed in the realm of ML. It operates by identifying an optimal hyperplane that maximizes the separation between data points from distinct classes within a high-dimensional space. SVM excels in situations where data isn't linearly separable, as it can utilize kernel functions to map the data into higher dimensions, enhancing separability. To facilitate nonlinear decision boundaries, the original feature space undergoes transformation into a new feature space.

The hyperplane function is denoted as

$$H(x) = \begin{cases} +1, if\ \omega.x + b \geq 1 \\ -1, if\ \omega.x + b \leq 1 \end{cases}$$

It is necessary to minimize the objective function so that $y_i(\omega.xi +)b \geq 1$ is always satisfied.

IV. Experimental Results And Discussion

This section describes the significance of the experimental data obtained from the IDS and displays the outcomes obtained. The ML and DL models were trained and tested using the Python libraries Scikit-learn and Keras.

A. Evaluation metrics

Statistical parameters play a crucial role in assessing the performance of models. The evaluation encompasses four commonly used metrics: true positive (TP), false positive (FP), and true negative (TN), false negative (FN), specifically employed for evaluating the effectiveness of IoMT attack categorization [XV].

Accuracy.

The percentage of accurate predictions. It can be determined by using:

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

Precision

The ratio of expected positive models.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{TN}}$$

Recall

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

F1- Score

$$\text{F1} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}$$

B. Ml Model Performance

In Figure 4, the performance of two machine learning techniques, namely SVM and RF, was evaluated for detecting IoMT attacks. It was observed that the RF technique outperformed the SVM models in the context of IoMT attack detection.

Table 2: Performance analysis of the proposed system

ML topology	Precision	Accuracy	F1 Score
RF	0.96	0.9982	0.961
SVM	0.953	0.951	0.947

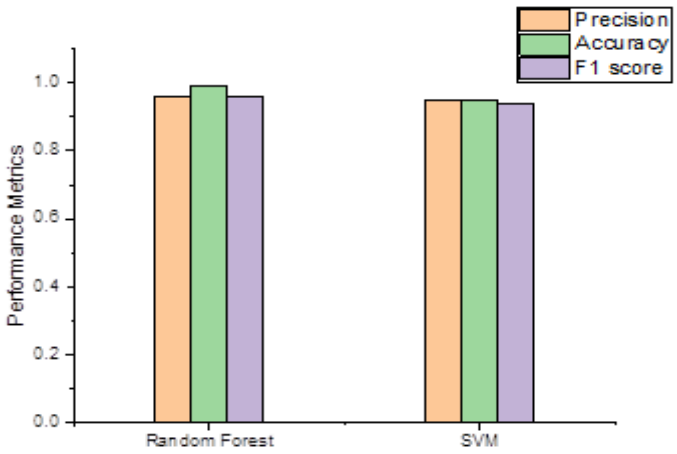

Fig 4. Performance analysis of the proposed ML topologies.

V. Comparative Analysis

Table 3 and Figure 5 depict the comparative analysis of the proposed system with the existing topology.

Table 3: Comparative Analysis.

Article	Techniques	Accuracy
Ref. [5]	KNN	92.06%
Ref. [4]	Tree Classifier	96%
Proposed Work	MWOA-RF	99.82%

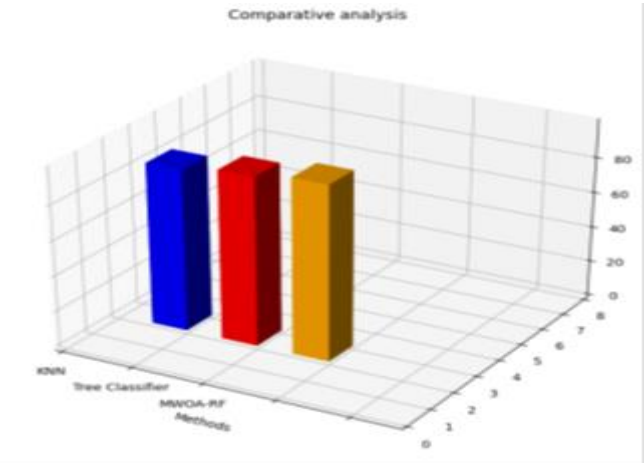

Fig 5. Comparative analysis

A K-Nearest Neighbors (KNN) methodology was introduced for IDS in IoMT-based networks [5]. The analysis revealed an impressive accuracy of approximately 92.06% for KNN. It's important to note that this approach did not involve feature extraction topology in the feature selection process.

However, in the study conducted by Gupta et al. [IV], efforts were made to enhance the performance of the IDS in IoMT by incorporating data augmentation and a tree classifier technique. Despite these improvements, this work demonstrated lower accuracy levels, particularly for real-time applications. Consequently, the proposed algorithm, MWOA in conjunction with the Random Forest (RF) model, has been introduced to elevate the performance and accuracy of IoMT-based IDS. The innovative approach demonstrated superior performance in comparison to existing state-of-the-art methodologies, presenting a promising solution to elevate intrusion detection capabilities within IoMT networks.

VI. Conclusion And Future Work

This study explores the design and analysis of an Intrusion Detection System (IDS) using machine learning (ML) within a smart healthcare system. Utilizing a dataset derived from a real-time testbed that includes both network traffic data and patient biometric information, the research focuses on enhancing IDS performance through the implementation of a Modified Whale Optimization Algorithm (MWOA) for feature selection.

The results of our analysis convincingly demonstrate the superior performance of the proposed MWOA-RF (Modified Whale Optimization Algorithm - Random Forest) approach compared to other ML algorithms used for IDS. The chosen performance metrics, which include accuracy, precision, recall, and F1-score, provide strong evidence of the effectiveness of the MWOA-RF approach in detecting intrusions. The study highlights the importance of feature selection in improving the accuracy and efficiency of IDS in smart healthcare systems. By leveraging advanced optimization techniques and robust ML algorithms, the proposed IDS framework offers a promising solution for enhancing security in IoMT environments.

Conflict of Interests:

The authors have no relevant financial or non-financial interests to disclose.

References

I. Awotunde Joseph Bamidele et al., : 'A deep learning-based intrusion detection technique for a secured IoMT system'. *International Conference on Informatics and Intelligent Applications*. Cham: Springer International Publishing, 2021. 10.1007/978-3-030-95630-1_4

II. Binbusayyis Adel et al., : 'An investigation and comparison of machine learning approaches for intrusion detection in IoMT network'. *The Journal of Supercomputing*. Vol. 78(15), pp. 17403-17422, 2022. 10.1007/s11227-022-04568-3

III. Ghubaish Ali et al., : 'Recent advances in the internet-of-medical-things (IoMT) systems security'. *IEEE Internet of Things Journal*. Vol. 8(11), pp. 8707-8718, 2020. 10.1109/JIOT.2020.3045653

IV. Gupta Karan et al., : 'A tree classifier based network intrusion detection model for Internet of Medical Things'. *Computers and Electrical Engineering*. Vol. 102, 108158, 2022. 10.1016/j.compeleceng.2022.108158

V. Hady Anar A. et al., : 'Intrusion detection system for healthcare systems using medical and network data: A comparison study'. *IEEE Access*. Vol. 8, pp. 106576-106584, 2020. 10.1109/ACCESS.2020.3000421

VI. Khan Soneila, and Adnan Akhunzada. : 'A hybrid DL-driven intelligent SDN-enabled malware detection framework for Internet of Medical Things (IoMT)'. *Computer Communications*. Vol. 170, pp. 209-216, 2021. 10.1016/j.comcom.2021.01.013

VII. Kumar P., Gupta G. P. and Tripathi R., : 'An ensemble learning and fog-cloud architecture-driven cyber-attack detection framework for IoMT networks'. *Computer Communications*. Vol. 166, pp.110-124, 2021. 10.1016/j.comcom.2020.12.003

VIII. Malamas Vangelis et al., : 'Risk assessment methodologies for the internet of medical things: A survey and comparative appraisal'. *IEEE Access*. Vol. 9, pp. 40049-40075, 2021. 10.1109/ACCESS.2021.3064682

IX. Manimurugan S., et al., : 'Effective attack detection in internet of medical things smart environment using a deep belief neural network'. *IEEE Access*. Vol. 8, pp. 77396-77404, 2020. 10.1109/ACCESS.2020.2986013

X. Nandy S., Adhikari M., Khan, M. A., Menon V.G. and Verma S., : 'An intrusion detection mechanism for secured IoMT framework based on swarm-neural network'. *IEEE Journal of Biomedical and Health Informatics*. Vol. 26(5), pp.1969-1976, 2021. 10.1109/JBHI.2021.3101686

XI. Radoglou-Grammatikis, Panagiotis et al., : 'A self-learning approach for detecting intrusions in healthcare systems'. *ICC 2021-IEEE International Conference on Communications*. IEEE, 2021. 10.1109/ICC42927.2021.9500354

XII. Ravi Vinayakumar et al., : 'A Multi-View attention-based deep learning framework for malware detection in smart healthcare systems'. *Computer Communications*. Vol. 195, pp. 73-81, 2022. 10.1016/j.comcom.2022.08.015

XIII. Rbah Yahya et al., : 'Machine learning and deep learning methods for intrusion detection systems in iomt: A survey'. *2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*. IEEE, 2022. 10.1109/IRASET52964.2022.9738218

XIV. Saheed Y. K. and Arowolo M. O., : 'Efficient cyber attack detection on the internet of medical things-smart environment based on deep recurrent neural network and machine learning algorithms'. *IEEE Access*. Vol. 9, pp.161546-161554, 2021. 10.1109/ACCESS.2021.3128837

XV. Unal Devrim, Shada Bennbaia, and Ferhat Ozgur Catak. : 'Machine learning for the security of healthcare systems based on Internet of Things and edge computing'. *Cybersecurity and Cognitive Science*. Academic Press, 2022. Pp. 299-320. 10.1016/B978-0-323-90570-1.00007-3