The Journal of Supercomputing (2022) 78:14866–14891
https://doi.org/10.1007/s11227-022-04453-z

Extreme learning machine and bayesian optimization-driven intelligent framework for IoMT cyber-attack detection

Janmenjoy Nayak[1] · Saroj K. Meher[2] · Alireza Souri [3] · Bighnaraj Naik[4] · S. Vimal[5]

Accepted: 14 March 2022 / Published online: 10 April 2022

Abstract
The Internet of Medical Things (IoMT) is a bionetwork of allied medical devices, sensors, wearable biosensor devices, etc. It is gradually reforming the healthcare industry by leveraging its capabilities to improve personalized healthcare services by enabling seamless communication of medical data. IoMT facilitates prompt emergency responses and provides improved quality of medical services with minimum cost. With the advancement of modern technology, progressively ubiquitous medical devices raise critical security and data privacy concerns through resource constraints and open connectivity. Vulnerabilities in IoMT devices allow unauthorized access for potential entry into healthcare and sensitive personal data. In addition, the patient may experience severe physical damage with the attack on IoMT devices. To provide security to IoMT devices and privacy to patient data, we have proposed a novel IoMT framework with the hybridization of Bayesian optimization and extreme learning machine (ELM). The proposed model derives encouraging performance with enhanced accuracy in decision-making process compared to similar state-of-the-art methods.

Keywords IoMT · Extreme learning machine · Bayesian optimization · IoT security

1 Introduction

The IoT (Internet of Things) is a conceptual term of device computing, which describes the idea of connecting physical objects to the internet every day and identifying themselves with other devices. IoT is a vital piece of technology with different things related to each other and is expected to develop dramatically over time

✉ Janmenjoy Nayak
jnayak@ieee.org

Extended author information available on the last page of the article

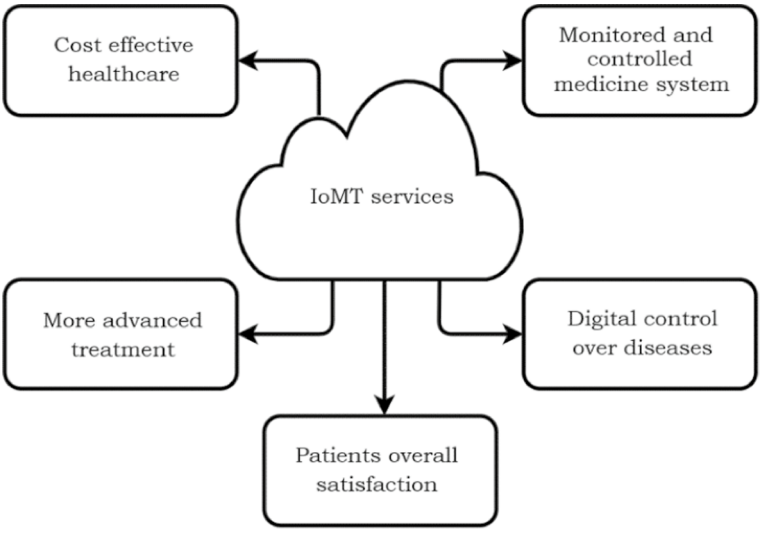


Fig. 1 IoMT services

for human well-being. There is untapped potential in the IoT, such as effective data-driven decision making, the ability to monitor and track things, greater workload ease with automation, increased efficiency and production by saving resources and money, and paving the way to a better quality of life. Rapid growth in communication and information technology with more excellent computing capabilities, leveraging the potential of IoT to the medical domain and named as the Internet of Medical Things (IoMT) technology. Healthcare IoT, or IoMT, states the connected infrastructure of medical devices and software applications that can communicate with various healthcare information technology systems [1].

IoMT itself is an umbrella term that encompasses different and varied types of medical devices employed at various stages of the healthcare process and in several unique capacities. The underlying concept behind all IoMT devices is that they capture, compile, and use data to streamline processes and maximize performance in various services involved in the medical sector (Fig. 1). IoMT systems use several smart actuators and biosensors, which are responsible for gathering real-time a nd confidential data related to patients. It allows medical professionals to understand and interpret sensitive information [2] better and more effectively. These sensors can be used as implants inside the human body and produce enormous real-time data for analysis and effective decision making. Over the next few decades, IoMT would explore the key challenges and trends in the world to transform the healthcare sector with its great potential for many applications, from remote monitoring to the incorporation of medical equipment. As a result, IoMT-based health care assessment can prevent fatal outcomes and increase the Nation's productivity. It results in bringing substantial growth for developing countries.

Moreover, IoMT can keep modern society functionally productive even in pandemic circumstances. IoMT with enabled devices unleashes the ability to maintain

the security and safety of patients and support doctors to provide excellent treatment. IoMT facilitates the continuity of healthcare services and keeps updating the patient status for those who need real-time and regular medical monitoring and preventive interventions [3]. IoMT also promotes the process of diagnosis and treatment, such as chronic disorders, exercise services, and elderly care. As interactions with doctors have become more effective and uncomplicated with increased patient involvement and satisfaction, remote patient health monitoring tends to avoid re-admissions and reduce hospital stays. Indeed, it introduces many benefits, as shown in Fig. 1, and more will be discovered as it continues to grow dramatically in the health sector.

The swift development of IoMT systems with advancements in data and device management leads to security vulnerabilities. Security-related issues, such as valid authentication, improper data transmission encryption, insecure interfaces, harmful firmware/software, and security concerns are significant apprehensions for any IoMT system. IoMT systems require an enhanced design framework that ropes in complexity management and scalability to avoid attacks. In addition to the challenges of control, the scopes and scales of IoMT devices give rise to another significant challenge, i.e., users' privacy. While attacks on IoT systems are getting progressively more common, security metrics frequently center on networks and software. The attackers can acquire control and carry out pernicious activities, e.g., assaulting other devices near the undermined or compromised nodes. The IoT-based healthcare system works with physiological conditions recorded by wearable sensors of the users. It can also be supplemented by contextual information to predict unusual patterns of the situation more precisely. The sensed data from various biosensors are timely communicated to the centralized storage (cloud) through ZigBee, Wi-Fi, or Bluetooth using smart devices. An attacker explores ways to exploit the vulnerabilities by scanning weak sensors and injecting malicious data to gain access and control the medical data. Unknowingly, the malicious data are updated into the centralized storage, as shown in Fig. 2.

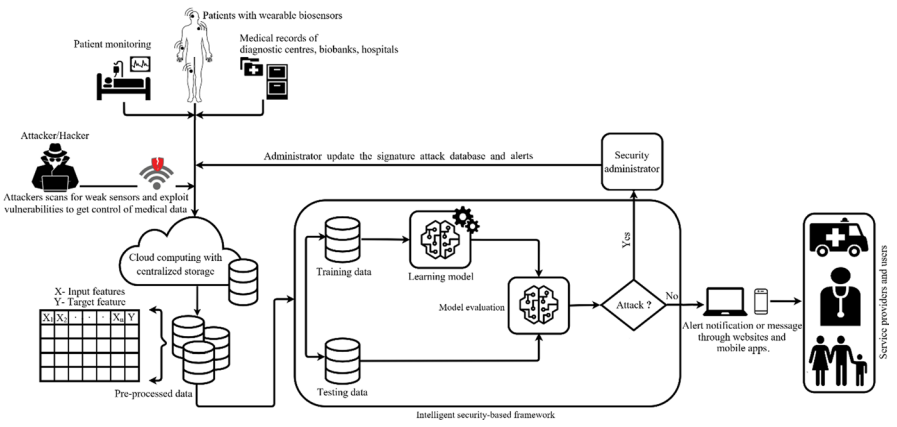


Fig. 2 General overview of the IoMT healthcare system

Mostly, IoMT devices do not have efficient malware or virus protection software. They are merely a reflection of the low-memory as well as low-power storing-based mechanism. The inaccessibility of malware and virus security on IoMT devices help to act as bots and transform the malicious activity to additional devices in the network. Moreover, to attack such devices, hackers can also access sensitive information gathered and transmitted through the IoMT devices. The absence of high security, integrity, and confidentiality of data in IoMT restricts its potential to disrupt the widespread implementation of this technology. With all the advantages, the IoMT application is accompanied by the possibility of vulnerabilities and new security breaches in the healthcare system. This is correlated with the following factors: (1) medical devices primarily capture and exchange confidential patient data, (2) IoMT design technology causes complexity and incompatibility issues, and (3) medical IoT system manufacturers do not prioritize security features. Major security concerns such as confidentiality, integrity, and availability are increasing due to the factors mentioned above. So, implementing apt security measures is very crucial.

As described above, we list the seven factors in Table 1, extending from security concerns to the threats of high client expectations. These factors of Table 1 are significant concerns for the development and growth of the IoT in the medical domain. The key to creating natural and long-term productivity and affluence through these incredible technologies will help overcome many issues.

To deal with such issues in the IoMT environment, several machine learning frameworks have been introduced in the recent past. Begli et al. [4] proposed a secure remote healthcare architecture system for remote monitoring (patients' data were collected) and responding appropriately in an emergency. Because of offering a safe framework against User to Root (U2R) and Denial of Service (DoS) attacks, a multi-layer (hybrid) intrusion detection system with a machine learning algorithm (support vector machine) was proposed. The proposed method is designed with security measures to deal with possible attacks that can minimize the uncertainty of complex decisions and allocate specific values to the outcomes of actions. However, because of the multi-agent-based layered and rule-based architecture, it consumes more energy. It incurs a substantial computational cost that makes the system ineffective in dealing with security measures of eavesdropping.

Marwan et al. [5] proposed machine learning techniques for preventing illegal access to medical records and personal information and provide secure data processing in a cloud environment. The offered machine learning (support vector machine) method was implemented using fuzzy C-means clustering to classify image pixels effectively. It also incorporated the module CloudSec for reducing the risk of the possible exposure of medical data through a conventional two-layered architecture. This approach aims to optimal feature extraction and could help cloud providers avoid expensive encryption methods for data protection. However, the model's consistency is not sure. This is because the classifier's performance may differ with the dataset types, which take more training time due to complex matrix operations of images and Gaussian assumptions that lead to higher computational costs.

After having a deeper perspective on these developments, it is observed that traditional machine learning algorithms are confronted with significant challenges in dealing with algorithm selection, data acquisition, time and resources, interpretation

Table 1 Challenges of IoMT in Healthcare

Challenge	Interpretations
Interoperability and compatibility of different IoMT devices	The continued interoperability and compatibility of IoMT devices are difficult. Because devices run different software versions, sensors with additional power consumption capabilities, various protocols, and security standards lead to security vulnerabilities and performance issues
Data overload with massive inputs of generated data	Accessing and processing large amounts of data creates problems for clinicians' decision making to handle complex medical issues
IoMT devices will increase the attack surface	IoMT healthcare devices allow attackers to gain unauthorized access to essential medical information by creating vulnerable security spots
Existing software infrastructure is obsolete	Information and communication technology infrastructures are not updated, which does not allow for the integration of new IoT devices
Connectivity and bandwidth issues	The IoT uses a centralized architecture with a server-client model to access the various system workstations and servers. But, iteratively, it fails if billions of devices use the same network, and the connectivity is a more significant challenge for the IoT applications to manage for space
Security	Malicious hackers do not need to break the plastic enclosure of an IoT device to access confidential information. Numerous security vulnerabilities are tied up with the IoT, so it is easy to finesse your way in due to devices with unpatched applications and other major security flaws
Government regulations	The common feature of technological advances should be of sudden interest for inclusion as government policy to keep up with the technology. With the rapid IoT development, the government is taking its time to keep up, and companies are often left without vital decision-making details

of results, and high error-susceptibility. Similarly, deep learning applications need large annotated data sets, and these are hard to obtain. Often, annotation is time-consuming, costly, and also vague. Additionally, over-trained NNs yield the worst generalization performance. Thus, validation and appropriate stopping criteria are required to minimize the cost function. The challenge of the backpropagation algorithm, i.e., local minima effect, makes it unstable. All these factors make the learning process unsuitable to use optimally.

Unlike the conventional machine learning models like NN and deep learning implementations, ELMs are single- or multi-hidden-layer feed-forward neural networks used to solve much real-world data mining and other complex problems. In ELM, the weight parameters between the input and the hidden layers need not be tuned, and it is a straightforward computation approach. The number of hidden nodes is arbitrarily assigned and has an automatic updating procedure with the

corresponding ancestors. The experimental results provide more efficient generalization performance with high-speed learning than classical, popular learning strategies for classification-based problems and approximation for objective benchmark functions. In the traditional approach of tuning the hyperparameters using grid search and randomized search, it is required to evaluate each set of hyperparameters by using the objective function, which becomes cost computing while dealing with a large no. of hyperparameter settings. Moreover, it becomes worst when the training time of the model is high.

Keeping all these aspects in view, a novel hybrid Bayesian optimization and ELM-based lightweight framework are designed to identify malicious access in the IoMT environment. The major contributions of this article are as follows:

i. An optimized ELM model has been designed to identify and mitigate malicious activities in the IoMT environment, using an efficient Bayesian optimization approach. The method has been adopted for finding the optimal set of hyperparameters of ELM to analyze the bigdata as a part of sensors and IoT devices in an IoMT environment.
ii. An intelligent hybrid security framework is designed by using a realistic dataset named ToN_IoT [23] to realize the impact of security measures on a dynamic scenario.
iii. The performance evaluation of the proposed model is made with the state-of-the-art ensemble and other conventional machine learning-based methods to realize its efficacy over others.

The rest of the paper is organized as follows. Section 2 discusses the literature study of the related IoT security research with intelligent methods. Section 3 elaborates the proposed Bayesian optimized hyper-tuned ELM approach to detect intrusive activities in an IoMT environment. The environmental setup for this experiment and the analysis of the results are described in Sect. 4. In this section, a rigorous performance analysis among all the machine learning, ensemble learning-based methods, and the proposed method is conducted to analyze the proposed method's efficiency compared to others. Section 5 concludes this work with a few critical future concerns.

2 Literature study

Several types of research have been carried out for dealing with the security issues of IOMT network-based devices. Newaz et al. [6] proposed a smart healthcare system that used implantable wearable and medical devices to monitor patients' vital signs continuously and robotically detect them to prevent critical medical conditions. A machine learning framework with a security-based health guard was proposed to detect malicious activities. The health guard perceives the different vital signs of connected devices and correlates them to understand the functionalities of the patient's body to distinguish between normal and malicious activities. Machine

learning techniques such as artificial neural network (NN), decision tree (DT), random forest (RF), *k*-nearest neighbor (KNN) are applied to detect malicious activities in a smart healthcare system. The computation of the proposed model is easy, and accuracy can be improved for large datasets. But, the results cannot be reliable in some instances when dependence (or correlation) exists between variables. Also, it maximizes the ambiguity of complex decisions and does not assign precise values to the outcomes of actions when data is limited.

He et al. [7] proposed connected healthcare systems for remote monitoring of patients' physical conditions. The paper deals with the security aspects and vulnerabilities of the systems and derives a new intrusion detection method based on a stacked autoencoder. The central part of the connected healthcare system is composed of three parts, i.e., the acquisition unit for human physiological data, the field control unit, and the remote monitoring center, client–server architecture. The major advantage of the proposed model is to reduce the feature dimensions through extracting more distinguishing features and can detect derivative attacks that have not occurred earlier. However, it is difficult to interpret the NN's processing elements called the BlackBox phenomenon and requires high processing time, making it unsuitable for large-coupled data.

Al-Shaher et al. [8] noted that the recent development of malicious codes had raised significant security concerns for patients' unauthorized access to electronic health records. The author(s) proposed an intelligent healthcare security system that includes a wavelet neural network approach, a smart firewall, an intelligent network intrusion detection subsystem, and an intelligent web filter for dealing with unwanted security threats. Multi-layer perceptron NN is also used to detect and classify attack analysis mechanisms. The analyzed structures of wavelet neural networks are used to develop an optimal neural network paradigm for the security problem. The projected model minimizes the ambiguity of complex decisions and assigns precise values as an outcome of actions. It can process high-dimensional data. The model is highly scalable and self-organized in detecting attacks. But, it demands high computational costs due to packet-based classification and probabilistic graphical rules.

To implement healthcare applications in a distributed environment, Lakhan et al. [9] have analyzed various offloading and scheduling problems in IoMT fog-cloud network. Further, the authors have developed a novel framework based on deep reinforcement learning and blockchain-enabled approaches. The framework consists of multi-criteria offloading that makes use of policies of deep reinforcement learning and blockchain-enabled task scheduling algorithms including task sequencing for the implementation of healthcare applications in distributed IoMT environment. The empirical results reveal that the suggested deep reinforcement learning and blockchain-enabled approaches enhance the performance of the framework by minimizing the computation cost and communication time in the distributed environment.

Lakhan et al. [10] designed a novel and cost-effective IoMT architecture based on blockchain-enabled fog cloud technology for minimizing the cost of healthcare applications and to provide security to the data in the healthcare networks. To minimize the cost, the framework makes use of the blockchain-enable smart-contract cost-efficient scheduling algorithm framework (BECSAF) scheme. The framework

makes use of the smart-contract blockchain scheme to provide data consistency and symmetric cryptography algorithm for validation. Moreover, experimental results show that the suggested algorithm schemes obtain better performance in terms of implementation of healthcare applications when compared with the standard approaches.

To implement changes in the dynamic environment, a novel security framework has been developed by Lakhan et al. [11]. The proposed framework makes use of the deep neural networks energy cost-efficient partitioning and task scheduling (DNNECTS) algorithm. The framework consists of the components namely application partitioning, task sequencing, and scheduling for processing critical healthcare tasks in the dynamic experiment. Moreover, the empirical results indicate that the suggested framework outperforms in the dynamic environment in terms of applications' cost and energy utilization.

To efficiently manage the resources, a healthcare resource management optimization(HRMO) framework has been suggested by Mutlag et al. [12]. In the suggested framework, fog computing has been incorporated as an intermediate layer to reduce the drawbacks of cloud computing. The chain fog nodes in fog computing are used to process the healthcare crucial tasks through the utilization of the MAS (multi-agent system) which is considered as the major responsibility of fog computing. Thus, MAS plays a major role in connecting all processing levels, namely edge, fog, and cloud in the proposed framework.

Golec et al. [13] have developed a security and privacy-based lightweight architecture known as iFaaSBus to protect the data acquired from IoT devices and to forecast the trend of the ailment. To diagnose the COVID-19 infection and to efficiently manage the resources, the developed framework makes use of machine learning, IoT, Function as a Service (FaaS). The patient's health data are secured using the OAuth-2.0 Authorization protocol-based privacy and JSON web token and transport layer socket (TLS) protocol-based security provided by the iFaaSBus. Further, the model has been validated using various machine learning approaches. It was evident from the results that KNN (K-nearest neighbor) attained better performance with an accuracy of 97.51% when compared with other approaches. It is also evident from the results that iFaaSBus attained better performance in comparison to non-serverless computing in terms of response time.

A model based on machine learning has been proposed by Yuvaraj et al. [14], for the parallelization of the jobs allocated and minimization of the runtime problems of the serverless frameworks. The suggested approach makes use of the GWO (gray wolf optimization) approach to enhance the mechanism of task allocation. In addition, the suggested approach also makes use of reinforcement learning (RIL) to optimize the GWO parameters which in turn enhances the task allocation mechanism. The simulation outcomes indicate that the suggested GWO-RIL approach provides reduced runtimes and accustoms with differing load conditions.

To easily manage relative 3D distances, a centralized heterogeneous formation flight position control design based on LQR PI (linear quadratic regulator proportional integral) controller has been proposed by Pirbhulal et al. [15]. In the proposed model, two wingmen quadcopters are used to track the output of the leader quadcopter. The pole placement control method and LQR PI control methods are used

to control the leader and the two followers, respectively. During the flight, formation geometry may be alternated to arbitrary shape using a control scheme if it is incorporated with collision avoidance mechanism. Further, singular values are used to analyze the closed-loop system stability. Finally, the proposed approach has been validated using MATLAB/Simulink and the results indicate that the model attains promising results even in the existence of critical perturbations in terms of output tracking and stability of the leader.

To efficiently classify the attacks by the intruders in the IoMT environment, a hybrid approach known as PCA-GWO (principal component analysis-grey-wolf optimization) based on the deep neural network has been suggested by Priya et al. [16]. Initially, categorical data are transformed into numerical data by utilizing a one-hot encoding scheme. Then, PCA and GWO are applied to the pre-processed dataset to minimize the attribute dimensions to choose highly significant attributes. Further, various machine learning approaches known as Naïve–Bayes, support vector machine, K-nearest neighbor, random forest, and deep neural networks have been used for the classification of the reduced dataset. The experimental results indicate that the proposed approach obtains better performance with a 15% enhancement in accuracy and a 32% reduction in time complexity when compared with the conventional machine learning approaches in the efficient classification and prediction of cyber-attacks in the IoMT environment.

To predict the network resource consumption and to enhance the transmission of IoT services on time, a model based on machine learning and SDN (software-defined network) framework has been presented by Haseeb et al. [17]. The proposed framework makes use of the SDN centralized model to minimize the overhead caused by the control plane in the deployed network of IoT. In addition, the proposed approach makes use of a machine learning approach to optimize the performance of routing in a real-time environment. Then, the proposed approach makes use of dynamic metrics and SDN architecture for the prediction of link status and refinement of the strategies. Finally, the SDN controller makes use of a security algorithm for the efficient management and safeguard of the IoT nodes from anonymous occurrences. From the experimental outcomes, it was observed that the developed model obtained better performance in terms of network throughput and data delay by 10% and 21%, respectively.

For accurate identification of the brain tumor concerning its grade at the early stage, an automated security system that makes use of PART (partial tree) has been presented by Khan et al. [18]. Further, the proposed approach has been validated using tenfold cross-validation and an advanced feature set that has not been utilized formerly for the accurate recognition of the brain tumor. From the empirical results, it was identified that the suggested approach obtained better performance in terms of computational cost and accuracy when compared with other approaches such as random tree, Naive Bayes, rep tree, and random forest.

For the diagnosis of age-related macular degeneration (AMD) disease, a teleophthalmology framework based on scalable cloud technology that makes use Internet of Medical Things (IoMT) has been proposed by Das et al. [19]. The proposed framework forwards the retinal fundus images that were captured from the head-mounted camera of the patients to their personal and secure cloud storage for the

prediction and detection of the severity of the AMD disease. Further, the severity of AMD disease is detected and identified by analyzing the images using the AMD-ResNet convolution neural network which makes use of 152 layers. For diagnosing the disease severity, the proposed model was trained using 130,000 AREDS (age-related eye disease study) fundus images acquired over 12 years from the NIH (National Institute of Health). From the experimental results, it was observed that the proposed model obtained $94.97 \pm 0.5\%$ sensitivity and $98.32 \pm 0.1\%$ specificity, respectively. Moreover, the proposed framework also makes use of temporal long-short term memory (LSTM) deep neural network for the prediction of advancement of AMD disease and precision medicine.

3 Proposed system

In this work, an ELM-based model [20, 21] with optimized parameters is developed for the efficient detection of intrusive behaviors in an IoT framework (Fig. 3).

The proposed problem can be visualized as an optimization problem where the objective is to select the best $p_i = \{f_i, H_i, \alpha_i\}$ in $P = \{p_1, p_2 ... p_n\}$ (population with 'n' number of hyper-parameter sets). Here, p_i is the ith randomly generated hyper-parameter value set which is drawn an allowed range of values as follows: $f_i \in \text{list}[1, 2, 3, 4, 5, 6, 7, 8]$, $H_i \in \text{range}[1, 60]$, and $\alpha_i \in \text{range}[0.1, 1.0]$. Here f_i ,H_i, and α_i are ith activation function, the selected number of the hidden layer, and the learning rate. The activation function f_i is chosen as '1,' '2,' '3,' '4,' '5,' '6,' '7,' and

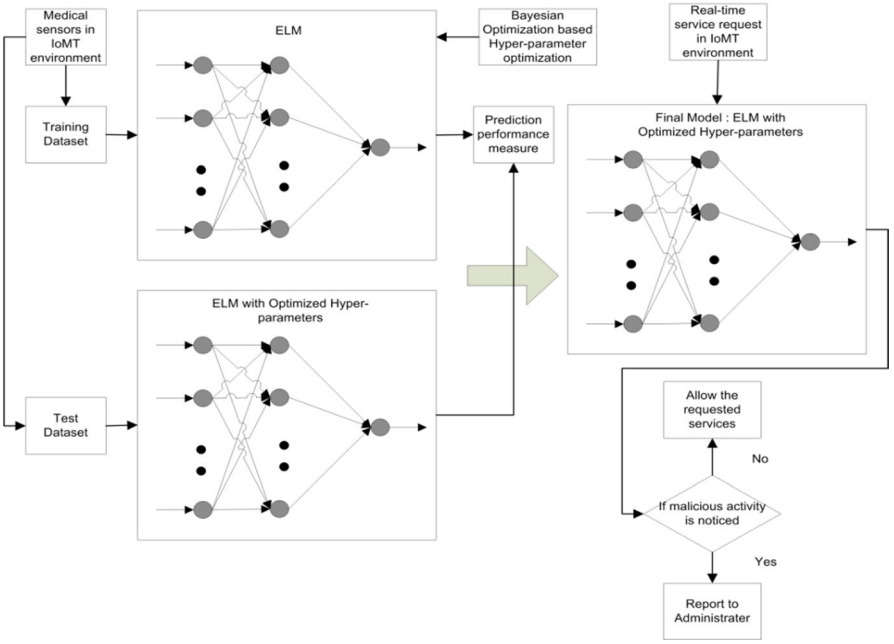


Fig. 3 Proposed system architecture

'8' for *Sine*, *Tanh*, *Tribas*, *Sigmoid*, *Hand* lim,*Soft* lim, *Gaussian*, and *Multiquadric*, respectively. The performance of ELM on the prediction of attack type is dependent on these parameters f,H, and α. On the given baseline ELM model, the impact of α ,H, and f on the prediction performance is shown in Figs. 4, 5, and 6, respectively. Here, the studied problem can be visualized as an optimization problem to get optimal $p_i^* = \{f_i, H_i, \alpha_i\}$ in P, which is the optimal parameter set of ELM for solving identification various attack types in IoT network. On the given IoT accesses profiles with connection traces $X = \{x_i, y_i\}_{i=1}^m$, $P = \{p_i\}_{i=1}^n$ and model $ELM(p_i, X)$, here the objective is to find optimal p_i^* which optimize the following objective function (Eq. 1):

$$\begin{aligned} p_i^* &= \arg\max_{p_i \in P} \{s_i = \text{score}(y, \hat{y} = ELM(p_i, X))\} \\ &= \arg\max_{p_i \in P} \left\{ s_i = \text{score}(y, \hat{y}) = \frac{1}{m} \sum_{i=1}^{m} I(\hat{y}_i, y_i) \right\} \end{aligned} \tag{1}$$

The connection traces dataset $X = \{x_i, y_i\}_{i=1}^m$ is the collection of instances x_i with 42 features and one class label at_i representing seven different attack types and one normal type. The proposed ELM model has been trained with these instances. In ELM, the prediction of the class label is made using Eq. (2), where H_out(Eq. 3) is the output matrix, β (Eq. 4) is the weight matrix representing the weights between hidden layer neurons and out neuron, and (Eq. 5) is the prediction.

$$\hat{y} = H\_\text{out} \times \beta \tag{2}$$

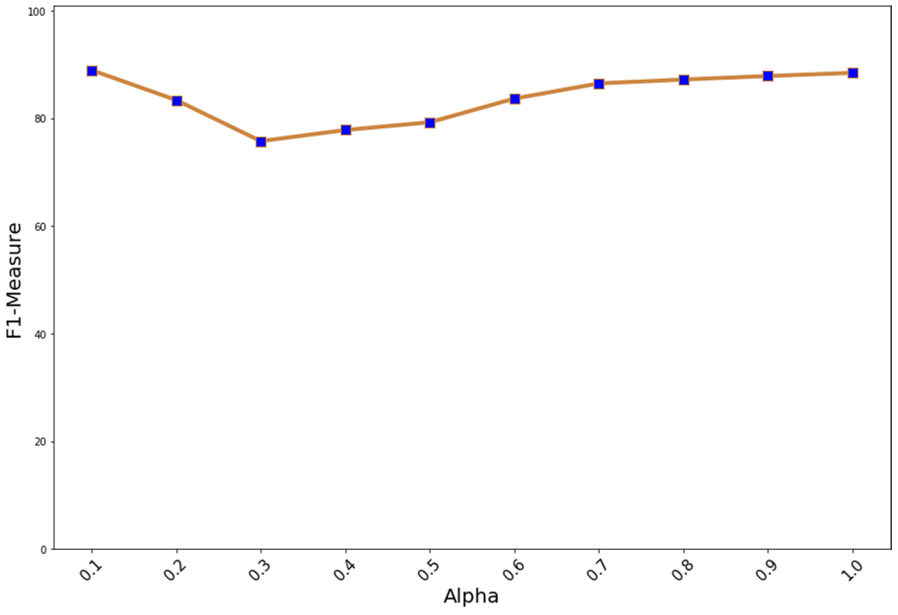


Fig. 4 Study on the impact of learning rate (α) on F1-measure

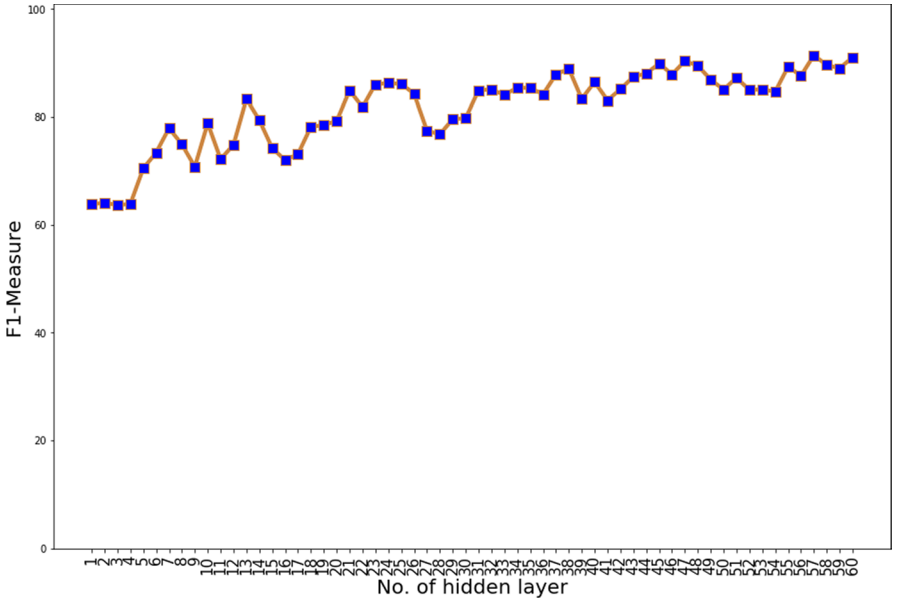


Fig. 5 Study on impact of no. of hidden layer (H) on F1-measure

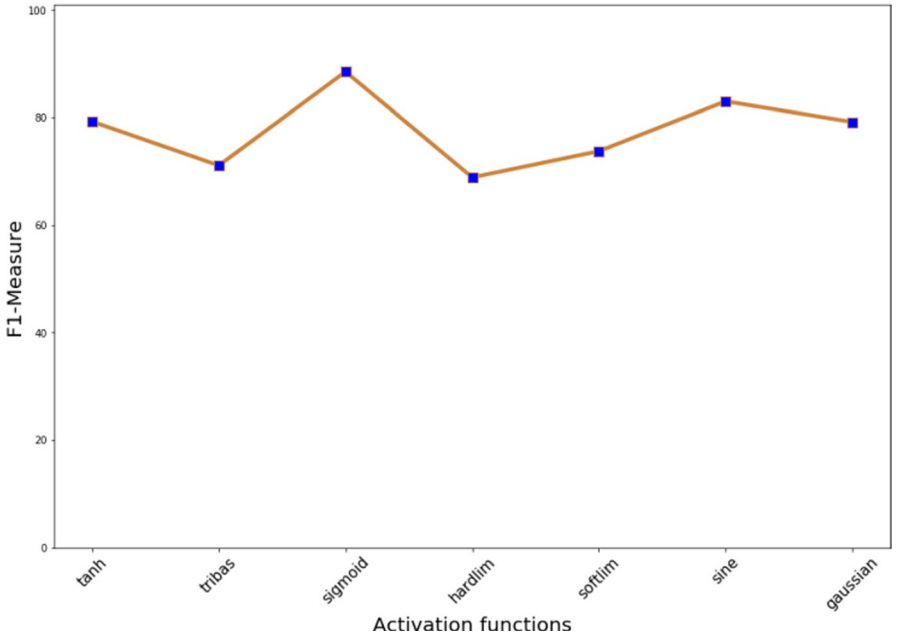


Fig. 6 Study on the impact of activation functions (f) on F1-measure

$$H\_\text{out} = \begin{bmatrix} f\left(b_1 + x_1 \times w_1\right) & \ldots & f\left(b_L + x_1 \times w\right) \\ \ldots & \ldots & \ldots \\ f\left(b_1 + x_N \times w_1\right) & \ldots & f\left(b_L + x_N \times w\right) \end{bmatrix}_{N \times L} \tag{3}$$

$$\beta = \left[\beta_1, \beta_2 \ldots \beta_L\right]^{\mathrm{T}}_{L \times 1} \tag{4}$$

$$\hat{y} = \left[y_1, y_2 \ldots, y_N\right]^{\mathrm{T}}_{N \times 1} \tag{5}$$

In this work, the hyperparameters of ELM are optimized with Bayesian optimization [22]. For unlabeled and complex (big) size data, this optimization often performs well (especially for unprecedented functions) to optimize the objective function. In this considered problem, the objective is to find out the optimal set of hyperparameters that maximize the score function defined in the objective function (Eq. 1). The ELM is successful and widely adopted for ease of implementation, incremental learning, batch learning, and sequential learning due to its efficiency and learning speed, generalization ability, and fast convergence. It is different from a traditional neural network it makes use of the Moore–Penrose generalized inverse technique for weight adjustment.

Bayesian optimization is an efficient choice over the grid and randomized search for optimizing hyperparameters as it searches the hyperparameters' values in the search space in an informed manner. Bayesian optimization initiates the search in hyperparameter search space from a small region of interest by using a surrogate function which is an approximation to the used objective function. Initial sample candidate solutions (points) in the search space are selected, and the surrogate function is obtained. Then, the obtained surrogate function is used to identify potential candidate solutions. Based on identified potential candidate solutions, the surrogate function is updated and other promising regions are identified. In each iteration, identify and focus on more regions of interest by updating the surrogate function. Unlike other optimization approaches that optimize an objective function, the Bayesian optimization technique optimizes the surrogate function. It makes use of a probabilistic model that computes the probability of the score on a given set of hyperparameters' values (surrogate function). These scores are used to select suitable hyperparameters' values and to guide the search. It usually provides better searching time as it focuses on those areas in the search space having a better probability of score. Bayesian optimization operates along with probability distribution for each hyper-parameter that it will sample from. These distributions have to be set by the users. Initially, this optimization approach starts with a wide search space and gradually focuses on a specific area around the optimal parameter retrieved from previous iterations. The domains of the hyper-parameters are defined along with many distribution functions.

In this work, the domains of the considered hyperparameters are as follows: activation function (f), alpha (α), and several hidden neurons (H). Here, the f is the mathematical equations that are responsible for determining whether the neuron input is significant for prediction. The α is the controlling parameter for the

adjustment of weights and H is the number of hidden neurons that highly impact the performance and network stability.

This work is focused on the process of finding optimal parameters (p_i) of ELM by using Bayesian optimization. The main components of Bayesian optimization are the objective function, surrogate function, and selection function. The objective function is used to evaluate the parameters combinations (p_i) and output a score $s_i = ELM(p_i, X)$ which indicates how well the set of hyperparameters performs for the considered problem. For the present problem, we have considered 'F1 score' as the evaluation matrix, and it is the objective to maximize the objective function presented in Eq. (1). Here, the objective is to put the restriction on the number of evaluations of objective functions.

Algorithm – 1: $s_i \leftarrow ELM(X, p_i)$

Let $X = \{x_1, x_2 ... x_N\}$ be the data, $x_i = \{x_{i,1}, x_{i,2} ... x_{i,N}, \text{at}_i\}$ where is the instance of the

1. Randomly generate bias b_i, $i = 0$ to L and weight w_i, $i = 0$ to L.

2. Calculate the hidden layer output function H_out by using selected activation function $f(.)$.

$$H\_out = \begin{bmatrix} f(b_1 + x_1 \times w_1) & \ldots & f(b_L + x_1 \times w) \\ \ldots & \ldots & \ldots \\ f(b_1 + x_N \times w_1) & \ldots & f(b_L + x_N \times w) \end{bmatrix}_{N \times L}$$

3. Compute the output weight matrix $\hat{\beta} = H\_out^{\dagger} \times y$, which maximizes the objective function $\left\| H\_out \times \hat{\beta} - y \right\| = \min_{\beta} \| H\_out \times \beta - y \|$. Here $H\_out^{\dagger}$ is the Moore-Penrose generalized inverse of the H_out.

$$H\_out^{\dagger} = (H\_out \times H\_out)^{-1} \times H\_out^{T}$$

Perform the prediction by using $\hat{\beta}$ on the data $\hat{y} = \alpha \times \left(H\_out \times \hat{\beta} \right)$, α is the learning rate.

$$\hat{y}_i^k = \begin{Bmatrix} 1 & if\ (y_i = k) \\ -1 & if\ (y_i \neq k) \end{Bmatrix}, \quad k = 1, 2, ..., \text{c}$$

Predict final class label as $\hat{y}_i^c = \arg\max_{k=1,2...c} \left(\hat{y}_i^k \right)$

Return score of the prediction $s_i = score\left(y, \hat{y} \right)$

Algorithm 1 takes a set of hyperparameters $p_i = \{f_i, H_i, \alpha_i\}$, train the model $ELM(p_i, X)$, and returns accuracy as a score. Here, we have considered negative score that our proposed optimization process requires a minimal value. Initially, n number of hyperparameters is generated randomly $P = \{p_i\}_{i=1}^{n}$ from the hyperparameter's domain distribution function, where each $p_i = \{f_i, H_i, \alpha_i\}$ represents ith hyperparameter set. Each p_i is evaluated by using the objective function $s_i \leftarrow ELM(X, p_i)$, which is meant for recognition of IoT security attack type. By this

process n number of pairs (p_i, s_i) are generated, from which the probability of set of hyperparameters given a score are broken down into two distributions as $l(P)$ and $g(P)$ as in Eq. (6).

$$\text{prob}(P| - s) = \left\{ \begin{array}{l} l(P) \;\; if - s < \text{threshold}(-s) \\ g(P) \;\; if - s \geq \text{threshold}(-s) \end{array} \right\} \tag{6}$$

By using a threshold λ on score, m (where $m < n$) number of hyperparameters $P' = \{p_1, p_2 ..., p_m\}$ are selected from P. Then, the expected improvements $E \times I_{s^*}(p_i)$ of each p_i in P' are computed to select the optimal hyperparameters set p_i^* to update the Surrogate function.

$$E \times I_{s^*}(P) = \frac{\lambda \times s^* \times l(P) - l(P) \times \int_{-\infty}^{s^*} \text{prob}(s)\text{d}s}{\lambda \times l(P) + (1 - \lambda) \times g(P)} \alpha \left(\lambda + \frac{g(P)}{l(P)}(1 - \lambda) \right)^{-1} \tag{7}$$

$$\text{prob}(s_i|p_i) = \frac{\text{prob}(p_i) \times \text{prob}(p_i|s_i)}{\text{prob}(s_i)} \tag{8}$$

Here, the objective is to maximize the $l(P)/g(P)$ to maximize the expected improvement. From P', the optimal p_i^* is selected which will maximize the Expected improvement $E \times I_{s^*}(p_i)$. By using the selected p_i^*, the Surrogate function is updated along with the feedback of the objective function (Algorithm 1). This process is repeated unless or until there is no change in optimal hyperparameters or no further improvement in expected improvement.

4 Experimental setup and result analysis

This section enlightens the experimental setup and evaluation of the proposed system based on the ensemble ELM approach and cloud architecture to mitigate the cyber-attack in a real-time IoMT environment.

4.1 Dataset collection and environment setup

In this study, the ToN_IoT dataset [23] that is obtained from a practical and large-scale network developed by UNSW Canberra Cyber IoT Lab, School of Engineering and Information Technology (SEIT), UNSW Canberra @ The Australian Defense Force Academy (ADFA) is considered for experimenting the proposed approach. The dataset consists of 43 features with a total of 4,61,043 observations, of which 3,00,000 are normal observations and 1,61,043 are cyber observations. We have performed the simulation using the machine with the following specifications. It is an HP (ProDesk 600 G2 MT) desktop with operating system: Windows 10 Pro 64-bit, Processor type: Intel(R) Core (TM) i7-6700 CPU with a capacity of 3.40 GHz (8 CPUs) and a memory of 4096 MB RAM for the experimentation. Further, software

packages of PYTHON library such as Pandas, Imblearn, and Numpy framework are utilized to analyze data better, a framework like Matplotlib is utilized for the visualization of the data, and sklearn and Mlxtend are utilized for the implementation of machine learning and ensemble-based stacking. In addition, the optimized parameters of all the deliberated models that have been considered for the optimization of distinct hyper-parameters of the ELM approach are illustrated in Table 2.

4.2 Result analysis

To prove the effectiveness of the proposed approach, distinct machine learning and ensemble learning approaches have been considered as a part of experimentation. Moreover, the effectiveness of the proposed approach is also evaluated by considering the k-fold cross-validation technique, which divides the dataset into a '*k*' number of folds in a random fashion. In this study, tenfold with stratified sampling is chosen to maintain efficient error estimation and less bias and variance. Further, the performance of all the considered approaches, along with the proposed approach, has been assessed by utilizing performance metrics such as precision, recall, F1 score, F2 score, Fbeta score, and ROC-AUC. In addition, particular focus has been given to the F1 score for accurately assessing the system's performance as the distribution of class labels is highly imbalanced and non-uniform.

In this study, a comprehensive relative analysis of the evaluation metrics utilized to assess all the ML and ensemble ELM approaches together with the proposed method is displayed in Table 3. The results show that the performance of the proposed method concerning the precision, recall, F1 score, F2 score, Fbeta score, and ROC-AUC values is superior when compared with the considered ML and ensemble ELM approaches. Figure 7a–g represents the actual versus predicted performance of all the deliberated ML approaches together with the proposed model on the activities of IoMT communication profiles.

From the results of the figures, it has been noticed that except for a few approaches like DT, NB, LR, and RF, all the other approaches such as XGBoost, ELM, and ELM along with Bayesian optimization have performed well. It is also noticed that the proposed approach surpasses all the considered approaches by achieving the highest predicted results in all the activities of the IoMT communication profiles.

Table 4 describes the experimental outcomes of the considered ensemble approaches such as ELM + RS, ELM + GA, and ELM + Bayesian optimization on stratified tenfold cross-validated data. From the empirical results represented in Table 4, it is realized that the proposed ELM, along with Bayesian optimization, has

Table 2 Optimized parameters of all considered models

Models	Optimized parameters
ELM + Bayesian Optimization	$f =$ 'sine,' $\alpha = 0.1$, $H = 60$
ELM + Genetic Algorithm	$f =$ 'sigmoid,' $\alpha = 0.1$, $H = 40$
ELM + Randomized Search	$f =$ 'sine,' $\alpha = 0.6$, $H = 20$

Table 3 Comparative analysis of performance metrics among all models

Prediction models	Performance metrics					
	Precision	Recall	F1 score	F2 score	Fbeta score	ROC-AUC
DT	0.769086	0.769086	0.769086	0.734934	0.692404	0.608390
NB	0.637984	0.637984	0.637984	0.572961	0.438794	0.5
LR	0.637984	0.637984	0.637984	0.572961	0.438794	0.5
RF	0.738423	0.738423	0.738423	0.695132	0.612288	0.595872
XGBoost	0.886420	0.886420	0.886420	0.869863	0.837399	0.692684
ELM	0.784730	0.784730	0.784730	0.764296	0.724570	0.636750
ELM + Randomized Search	0.748435	0.748435	0.748435	0.740861	0.728568	0.668570
ELM + Genetic Algorithm	0.978097	0.978097	0.978097	0.975687	0.971108	0.824434
Proposed Model (ELM + Bayesian Optimization)	0.990300	0.990300	0.990300	0.989175	0.986652	0.870034

surpassed the other two approaches in almost all the folds of stratified tenfold cross-validated data.

The finding of the AUC-ROC curve of the ELM approach, along with the proposed ELM and Bayesian optimization approach, is depicted in Fig. 8a, b. It is noticed from the visualization findings shown in Fig. 8a, b that both the ELM and ELM along with Bayesian optimization have displayed unremitting performance in every class of the IoMT communication profiles. Moreover, the actual versus predicted performance results using ensemble ELM approaches such as ELM + RS, ELM + GA, and ELM + Bayesian optimization are interpreted in Fig. 9a–c. From the visualization findings, it is identified that the predicted performance of ELM with Bayesian optimization has surpassed the other two approaches in all the classes of the IoMTcovisualisation profiles. The finding of the F1 score of all the deliberated ML approaches along with the proposed approach is represented in Fig. 10. It is identified from the figure that the proposed method has attained a higher F1 score when compared with the F1 score of the considered ML and other ensemble approaches. Therefore, from all the visual findings, it can be concluded that the proposed approach attains superior performance in all the classes of IoMT communication profiles and has proven to be the most robust approach for mitigating cyber-attacks in real-world applications of IoMT.

The performance of the proposed model is also compared with other existing similar research related to IoMT. Table 5 indicates the performance comparison of the proposed model with other similar intelligent models studied in the literature. The performance analysis shows the superiority of the projected method over other described methods to a larger extent. The experimental results show that the proposed system based on the cloud and ensemble learning approach has several advantages. The main advantage of the proposed method is that it can quickly identify malicious activities in highly dynamic and assorted networks of IoMT as the framework of the proposed system is simple and easy to implement. Another advantage is that few parameters are used in the training and testing phase to design the

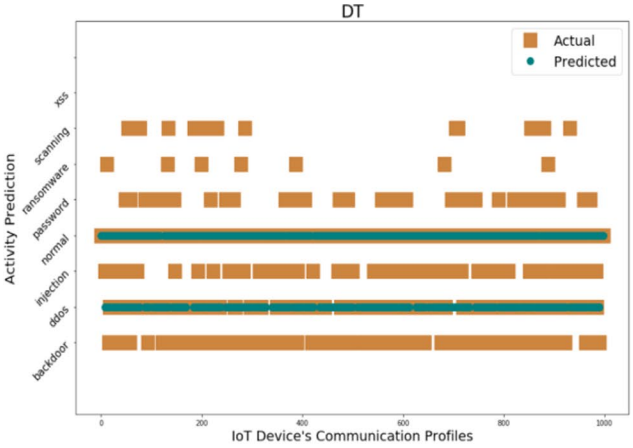
Fig. 7 AUC-ROC analysis on all the methods. (**a**) DT prediction performance. (**b**) NB prediction performance. (**c**) LR prediction performance. (**d**) RF prediction performance. (**e**) XGBoost prediction performance. (**f**) Prediction performance of ELM. (**g**) Prediction performance of ELM + Bayesian Optimization ▶

IDS. Moreover, these parameters can be easily updated in real-time predictions to enhance the overall efficiency of the proposed system regarding the accuracy, detection rate, low false-positive rate, and processing time.

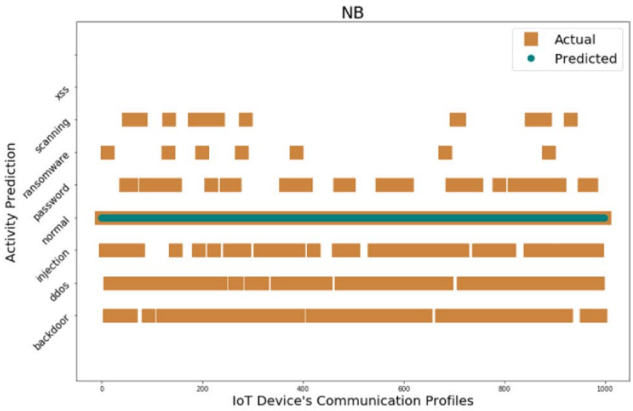
5 Conclusion and future work

In the last few decades, advanced technology and the population have increased day by day worldwide, which has increased the cost of healthcare and prices of services. Moreover, the rapid advancement in the IoT technology has contributed to vast development in the Internet of medical things (IoMT) which aims in enhancing the quality of patient's life. Therefore, the transformation of the healthcare sector to IoMT is required to assure a better quality of medical services at an affordable cost. The significant challenge in the deployment of the IoMT environment is the growing cyber-attacks. To protect the IoMT environment from unforeseen cyber-attacks, various researchers are functioning in this domain to provide different approaches for securing the IoMT environment from other cyber-attacks. This study proposed a hybrid approach based on ELM and Bayesian optimization that utilized cloud architecture to mitigate cyber-attacks in a real-time IoMT environment. Besides, the fundamental advantage of using an ensemble learning approach is to forecast advanced predictions by considering the predictions obtained from the individual considered machine learning approaches.

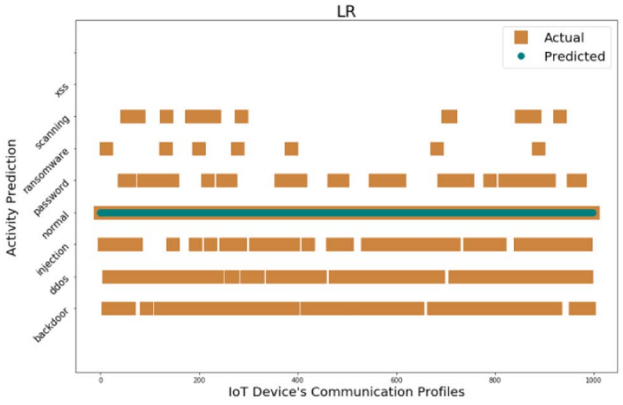
In this study, the effectiveness of the proposed ELM with Bayesian optimization for reducing cyber-attacks in the IoMT environment had been demonstrated by comparing the results with several machine learning approaches, such as DT, NB, LR, RF, XGBoost, ELM, and ensemble learning approaches, such as ELM with GA and ELM with RS. The experimental results demonstrated the superiority of the proposed method in terms of precision, recall, F1 score, F2 score, Fbeta score, and AUC-ROC curve with values 0.990300, 0.990300, 0.990300, 0.989175, 0.986652, and 0.870034, respectively, over the considered machine learning and ensemble methodologies. More focus may be attained in the future study of this work on the security and privacy concerns concerning multiple cloud/fog-based dynamic environments prolonged with extensive devices. Also, the adoption of ELM certainly works for few learned patterns, while it may fail for larger sized nonlinear data approximation based solutions. In this case, it is advisable to adopt deep learning neural network with better adaptability and learning capability for untrained features. Several security issues, such as hijacking, tampering messages, eaves dropping, device cloning, denial of service, denial of power attack [32], arise in the deployment of IoT technology as the components of IoMT are interconnected from different locations. In addition, deployment of IoMT technology also suffer from several limitations such as low battery



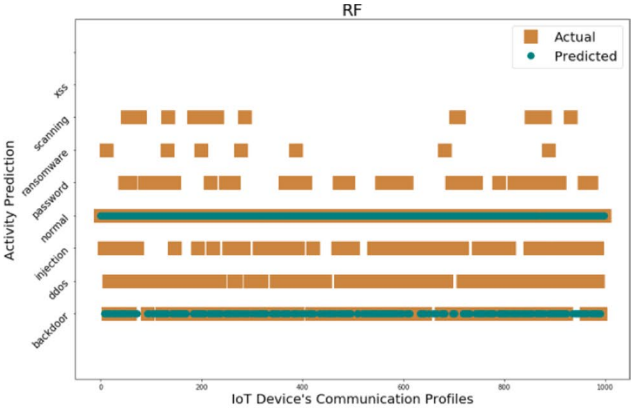
(a) DT Prediction Performance



(b) NB Prediction Performance

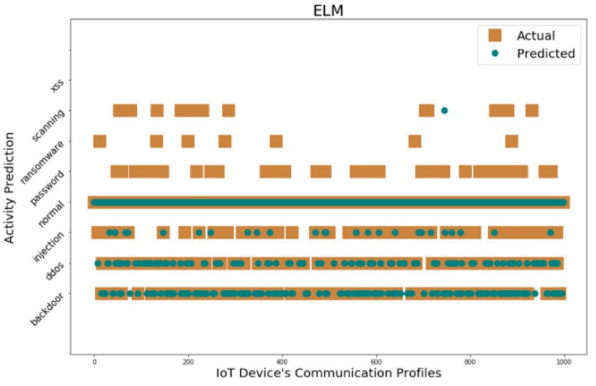


(c) LR Prediction Performance



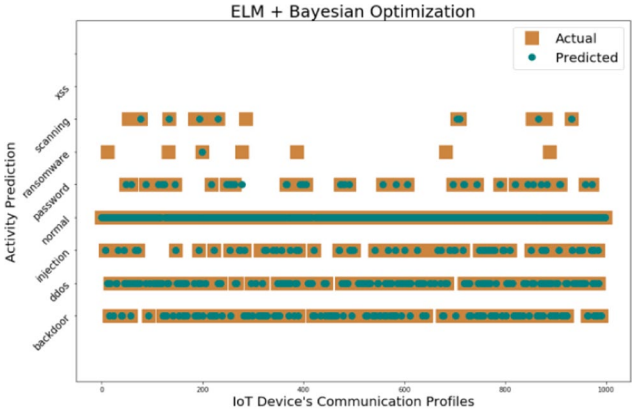
(d) RF Prediction Performance

(e)XGBoost Prediction Performance



(f) Prediction Performance of ELM

Fig. 7 (continued)



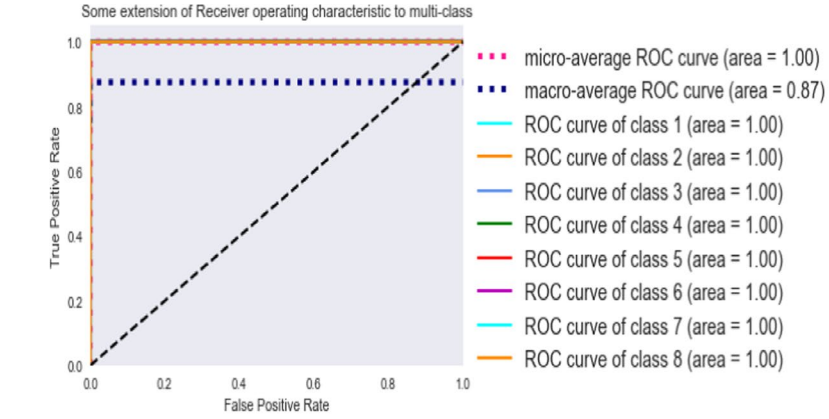
(g) Prediction Performance of ELM + Bayesian Optimization

Fig. 7 (continued)

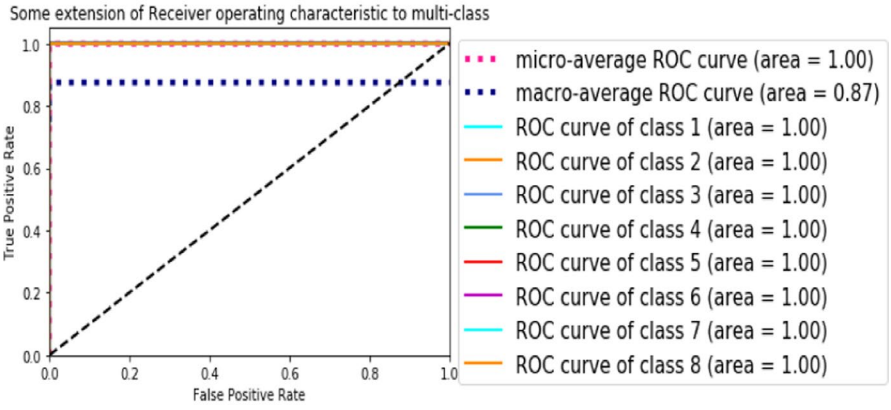
Table 4 Comparison of F1-measure among three ELM models

Stratified sampled cross-validated tenfold data	F1-measure		
	ELM + Randomized Search	ELM + Genetic Algorithm	ELM + Bayesian Optimization
Fold1	0.790625	0.959375	0.985625
Fold2	0.771875	0.9875	0.996875
Fold3	0.746875	0.978125	0.98125
Fold4	0.740625	0.978125	0.9975
Fold5	0.775	0.971875	0.978125
Fold6	0.75625	0.96875	0.965625
Fold7	0.774295	0.978056	0.987461
Fold8	0.793103	0.981191	0.987461
Fold9	0.724138	0.981191	0.994326
Fold10	0.742947	0.984326	0.994326

capacity, less processing power and less memory, interoperability and security. In this aspect, it is highly apt to adhere the intelligent methods to protect the devices and network. Moreover, based on the specific domain expert knowledge and other



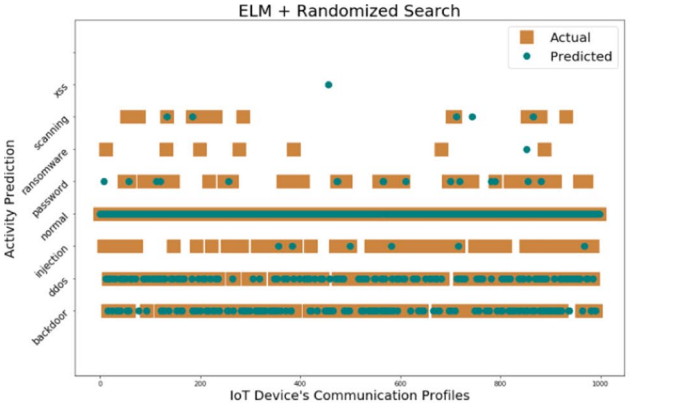
(a) ELM



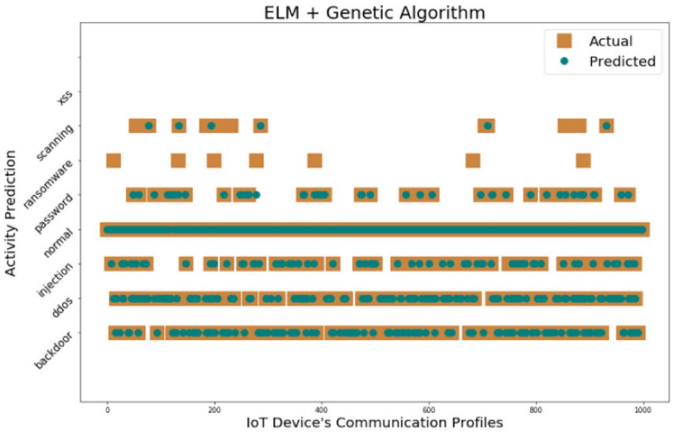
(b) ELM + BO

Fig. 8 AUC-ROC analysis of ELM and ELM + BO. (**a**) ELM. (**b**) ELM + BO

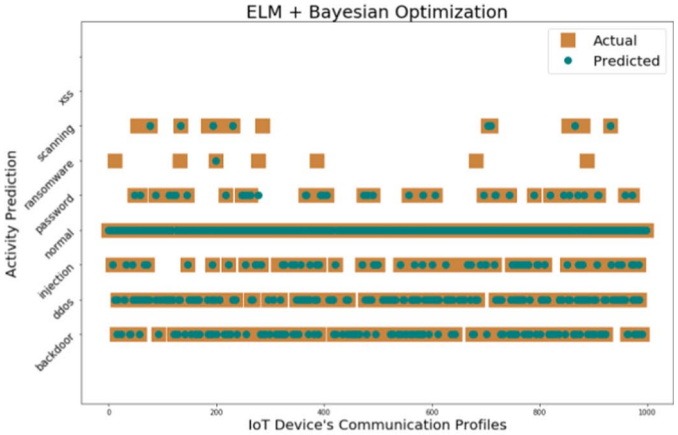
related information, the present study may be extended to ensure correct user-centric recommendations in a real-time cloud/fog-based environment.



(a)



(b)



(c)

Fig. 9 **a** ELM with randomized search, **b** ELM + Genetic Algorithm, **c** ELM + Bayesian Optimization

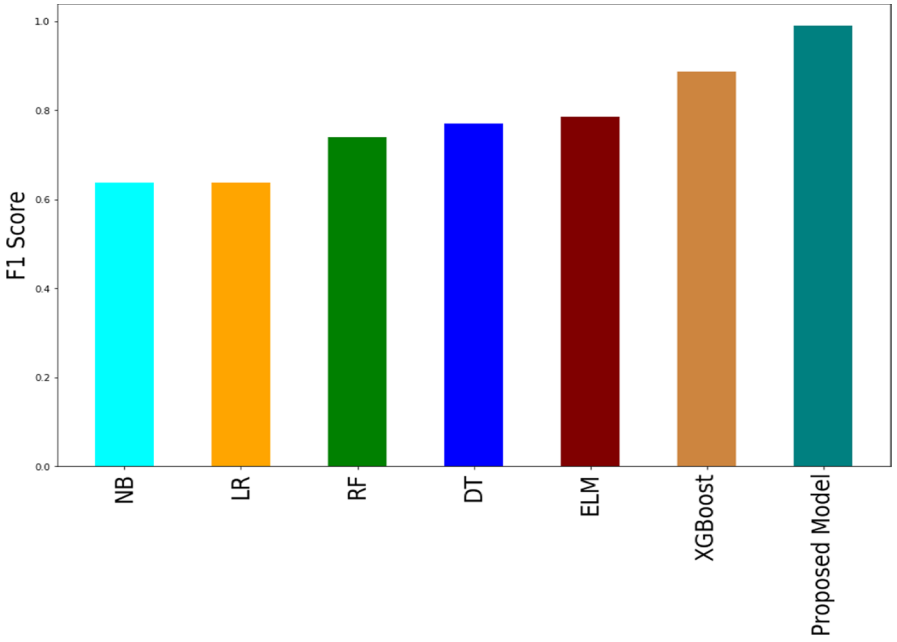


Fig. 10 F1 score comparison analysis of the studied models

Table 5 Performance comparison with other similar IoMT security models

Method name	Performance metrics	References
XGBoost	Accuracy of 96.35%	[24]
ML method (decision tree classifier, Naive Bayes, SVM, random forest, MLP, dense DNN Relu, Dense DNN Tanh)	Accuracy of 94.45%	[25]
Software-defined networking with reinforcement learning approach	Accuracy of 92.3%	[26]
Clustering with gray wolf model	Accuracy of 72%	[27]
Neural network-based voting System	Accuracy of 89%	[28]
Bagging, boosting, and voting ensemble	Accuracy of 93.2%, 92.7%, and 92.9%	[29]
Deep belief network	Avg accuracy of 97.73%	[30]
Hierarchical temporal memory	F1 score of 0.26	[31]

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Sayeed MA et al (2019) Neuro-detect: a machine learning-based fast and accurate seizure detection system in the IoMT. IEEE Trans Consumer Electron 65(3):359–368
2. Reddy DKK et al (2021) Exact greedy algorithm based split finding approach for intrusion detection in fog-enabled IoT environment. J Inform Secur Appl 60:102866
3. Ghubaish A et al (2020) Recent advances in the internet-of-medical-things (IoMT) systems security. IEEE Internet Things J 8(11):8707–18
4. Begli MR, Farnaz D, Hadis K (2019) A layered intrusion detection system for critical infrastructure using machine learning. In: 2019 IEEE 7th International Conference on Smart Energy Grid Engineering (SEGE). IEEE
5. Marwan M, Kartit A, Ouahmane H (2018) Security enhancement in healthcare cloud using machine learning. Proc Comput Sci 127:388–397
6. Newaz AKMI, et al (2019) Healthguard: a machine learning-based security framework for smart healthcare systems. In: 2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS). IEEE
7. He D et al (2019) Intrusion detection based on stacked autoencoder for connected healthcare systems. IEEE Netw 33(6):64–69
8. Al-Shaher MA, Hameed RT, Ţăpuş N (2017) Protect healthcare system based on intelligent techniques. In: 2017 4th International Conference on Control, Decision and Information Technologies (CoDIT), pp 0421–0426. IEEE
9. Lakhan A, et al. (2021) Mobile-fog-cloud assisted deep reinforcement learning and blockchain-enable IoMT system for healthcare workflows. Trans Emerging Telecommun Technol e4363
10. Lakhan A et al (2021) Smart-contract aware ethereum and client-fog-cloud healthcare system. Sensors 21(12):4093
11. Lakhan A, et al (2021) Deep neural network-based application partitioning and scheduling for hospitals and medical enterprises using IoT assisted mobile fog cloud. In: Enterprise Information Systems, pp 1–23
12. Mutlag AA, Ghani MK, Mohammed MA (2021) A healthcare resource management optimization framework for ECG biomedical sensors. In: Efficient Data Handling for Massive Internet of Medical Things, pp 229–244. Springer, Cham
13. Golec M, et al (2021) iFaaSBus: a security and privacy based lightweight framework for serverless computing using IoT and machine learning. IEEE Trans Ind Inform (2021)
14. Yuvaraj N, Karthikeyan T, Praghash K (2021) An improved task allocation scheme in Serverless computing using gray wolf optimization (GWO) based Reinforcement Learning (RIL) approach. Wireless Pers Commun 117(3):2403–2421
15. Pirbhulal S, et al (2019) Towards machine learning enabled security framework for IoT-based healthcare. In: 2019 13th International Conference on Sensing Technology (ICST). IEEE
16. Swarna Priya RM et al (2020) An effective feature engineering for DNN using hybrid PCA-GWO for intrusion detection in IoMT architecture. Comput. Commun. 160:139–149
17. Haseeb K et al (2021) A machine learning SDN-enabled big data model for IoMT systems. Electronics 10(18):2228
18. Khan SR et al (2020) IoMT-based computational approach for detecting brain tumor. Future Gener Comput Syst 109:360–367
19. Das A et al (2019) Distributed machine learning cloud teleophthalmology IoT for predicting AMD disease progression. Future Gener Comput Syst 93:486–498
20. Huang GB, Zhu QY, Siew CK (2006) Extreme learning machine: theory and applications. Neurocomputing 70:489–501
21. Huang GB, Zhou H, Ding X, Zhang R (2012) Extreme learning machine for regression and multiclass classification. IEEE Trans Syst Man Cybern Part B Cybern 42:513–529
22. Mockus J (2012) Bayesian approach to global optimization: theory and applications, vol 37. Springer, Berlin
23. Moustafa N (2019) ToN_IoT Datasets, 2019 (online). https://doi.org/10.21227/feszdm97. Accessed on 8th February 2021
24. Kumar P, Gupta GP, Tripathi R (2021) An ensemble learning and fog-cloud architecture-driven cyber-attack detection framework for IoMT networks. In: Computer Communications, vol 166. Elsevier, Amsterdam, pp 110–124. https://doi.org/10.1016/j.comcom.2020.12.003

25. Radoglou-Grammatikis P, Sarigiannidis P, et al (2021) A self-learning approach for detecting intrusions in healthcare systems. In: ICC 2021: IEEE International Conference on Communications. IEEE, pp 1–6. https://doi.org/10.1109/ICC42927.2021.9500354
26. Radoglou-Grammatikis P, Robolos K et al (2021) Modelling, detecting and mitigating threats against industrial healthcare systems: a combined SDN and reinforcement learning approach. IEEE Trans Ind Inform. https://doi.org/10.1109/TII.2021.3093905
27. Hatamian A, Tavakoli MB, Moradkhani M (2021) Improving the security and confidentiality in the internet of medical things based on edge computing using clustering. In: Computational Intelligence and Neuroscience.https://doi.org/10.1155/2021/6509982
28. Moukafih N, Orhanou G, El Hajji S (2020) Neural network-based voting system with high capacity and low computation for intrusion detection in SIEM/IDS systems. Secur Commun Netw 2020:1–15. https://doi.org/10.1155/2020/3512737
29. Saba T (2020) Intrusion detection in smart city hospitals using ensemble classifiers. In: 2020 13th International Conference on Developments in eSystems Engineering (DeSE). IEEE, pp 418–422. https://doi.org/10.1109/DeSE51703.2020.9450247
30. Manimurugan S et al (2020) Effective attack detection in internet of medical things smart environment using a deep belief neural network. IEEE Access 8(2):77396–77404. https://doi.org/10.1109/ACCESS.2020.2986013
31. Midani W, Fki Z, BenAyed M (2019) Online anomaly detection in ECG signal using hierarchical temporal memory. In: 2019 Fifth International Conference on Advances in Biomedical Engineering (ICABME). IEEE, pp 1–4. https://doi.org/10.1109/ICABME47164.2019.8940307
32. Gill SS, Tuli S, Xu M, Singh I, Singh KV, Lindsay D, Garraghan P (2019) Transformative effects of IoT, Blockchain and Artificial Intelligence on cloud computing: Evolution, vision, trends and open challenges. Internet Things 8:100118

Authors and Affiliations

Janmenjoy Nayak[1] · Saroj K. Meher[2] · Alireza Souri [3] · Bighnaraj Naik[4] · S. Vimal[5]

Saroj K. Meher
saroj.meher@isibang.ac.in

Alireza Souri
alirezasouri@halic.edu.tr

Bighnaraj Naik
bnaik_mca@vssut.ac.in

S. Vimal
svimalphd@gmail.com

[1] Department of Computer Science, Maharaja Sriram Chandra Bhanja Deo (MSCB) University, Baripada, Odisha 757003, India

[2] Systems Science and Informatics Unit, Indian Statistical Institute (ISI), Bangalore Centre, 8th Mile, Mysore Road, RVCE Post, Bangalore 560059, India

[3] Department of Computer Engineering, Haliç University, Istanbul, Turkey

[4] Department of Computer Application, Veer Surendra Sai University of Technology, Burla, Sambalpur, Odisha 768018, India

[5] Department of AI & DS, Ramco Institute of Technology, Rajapalayam, Tamil Nadu 626117, India