# Sine Cosine Algorithm with Optimal Convolutional Autoencoder for Intrusion Detection and Classification Model

**K. Hemavathi[1], Dr. R. Latha[2]**

[1]Research Scholar, Department of Computer Science, St. Peter's Institute of Higher Education and Research, Chennai
Email: *hemavathibhavani[at]gmail.com*

[2]Professor and Head, Department of Computer Science, St. Peter's Institute of Higher Education and Research, Chennai.
Email: *latharamavel[at]gmail.com*

**Abstract:** ***Network security comprises a multifaceted method that aims to protect computer networks from malicious activities, unauthorized access, and data breaches. The security mechanism is Intrusion Detection which is an important constituent that is employed to monitor and analyse the network traffic for recognizing and responding to intrusive or suspicious behavior. Innovative methods such as deep learning (DL) are employed to enhance the effectiveness of Intrusion Detection Systems (IDSs). DL is extremely implemented for IDS owing to its proficiency for automatically learning and extracting complex patterns and features from massive and multifaceted network datasets. Neural network (NN) models, permit the system to distinguish between anomalous patterns and normal network behaviors, increasing the accuracy of intrusion detection. The flexibility of DL methods to emerging cyberattacks with their adeptness to handle large - scale and various data, positions them as a strong and efficient tool for proactive and intelligent intrusion detection in existing cybersecurity settings. This article presents a Sine Cosine Algorithm with Optimal Convolutional Autoencoder for Intrusion Detection and Classification (SCAOCAE - IDC) method. The developed SCAOCAE - IDC system presents a wide - ranging strategy to improve the precision and effectiveness of IDSs. The method combines diverse advanced mechanisms like Min - Max scalar normalization for data preprocessing, Sine Cosine Algorithm (SCA) for feature selection (FS), Convolutional Autoencoder (CAE) for better feature extraction and classification, and Heap - Based Optimization (HBO) for hyperparameter tuning. The Min - Max scalar makes sure of robust data normalization, SCA increasingly chooses main features, CAE capably captures complex patterns in the data, and HBO fine - tunes hyperparameters for improved system performance. By employing the synergistic combination of such modules, the presented SCAOCAE - IDC algorithm indicates considerable outcomes for increasing the reliability and accuracy of IDSs and classification systems.***

**Keywords:** Intrusion Detection System; Cybersecurity; Sine Cosine Algorithm; Hyperparameter Tuning; Feature Selection

## 1. Introduction

With the extensive use of internet and improvements in accessible online content, cybercrime is also occurring at a higher rate. Intrusion detection is considered the primary stage to avoid security attacks [1]. Therefore, security methods like Intrusion Detection Systems (IDSs), Firewall, Intrusion Prevention Systems (IPS), and Unified Threat Modeling (UTM) have gained more attention in research. IDS identifies attacks from various techniques and network sources by gathering data followed by analyzing the data for potential security breaches [2]. The network based IDS analyzes the data packets that could be transmitted through the network and it is executed in two methods. Nowadays, anomaly based detection has been significantly behind the identification in which functions depend upon signature, and thus, anomaly based detection is a main domain for exploration [3]. The complexity of anomaly based IDS such as it requires handling new attacks without prior knowledge to identify the anomaly.

In recent times, research workers have considered that a more dependable IDS is embedded in the feature selection (FS) technique [4]. FS technique in major detection techniques has been employed for choosing fitted input features for base algorithms, with the objective of improving the recognition rate and reducing the false alarm rate (FAR) in Network - IDS (NIDS) [5]. Generally, the input features to classifiers can be always huge and not every feature is important to the classes to be categorized, therefore the requirement for an FS method [6]. However, FS technique is basically categorized into three various techniques embedded method, wrapper technique, and filter method [7]. The filter technique is mutual to majority of the FS algorithms that reliant on the choice of the fittest features under the dataset statistics without regard to the effectiveness of the method [8]. Alternatively, the wrapper method is more suitable, because the classifier efficiency will be employed in the exploration for the fitness assessment of the feature subsets leading to higher classification accuracy [9]. Numerous anomaly based methods are developed comprising Decision Tree (DT), Support Vector Machines (SVM), Linear Regression (LR), k - nearest neighbor (KNN) technique, Naive Bayes (NB) classifier, Genetic Algorithm (GA), Gaussian mixture model [10].

This article presents a Sine Cosine Algorithm with an Optimal Convolutional Autoencoder for Intrusion Detection and Classification (SCAOCAE - IDC) method. The method combines diverse advanced mechanisms like Min - Max scalar normalization for data preprocessing, Sine Cosine Algorithm (SCA) for feature selection (FS), Convolutional Autoencoder (CAE) for better feature extraction and classification, and Heap - Based Optimization (HBO) for hyperparameter tuning. The Min - Max scalar makes sure of robust data normalization, SCA increasingly chooses main features, CAE capably captures complex patterns in the data, and HBO fine - tunes hyperparameters for improved system performance. By employing the synergistic combination of such modules, the presented SCAOCAE - IDC algorithm indicates considerable outcomes for increasing the reliability and accuracy of IDSs and classification systems.

## 2. Literature Survey

Ramu et al. [11] presented new DL - based techniques and meta - heuristic optimization to increase the effectiveness of NIDS. Next pre - processing, an extended Pelican Optimization method (Ex - Pel) was deployed to choose the group of features in the preprocessed data in optimum way. In conclusion, the Self - Attention Assisted Weighted - AE (SAttn_WAE) method has been performed to identify the attacks exactly by using the sets of optimum features. In [12], the cyber - physical systems (CPS) environment was introduced as a layered technique like application layer, network layer, and physical layer as regards functionality. Afterward, various CPS attacks with regard to every layer have been expanded. Moreover, complexities and main problems related to all the layers. Subsequently, deep learning (DL) techniques have been evaluated for malicious URLs and intrusion detection in CPSs. In [13], a convolutional recurrent neural network (CRNN) was utilized for making a DL - based hybrid IDS that identifies the attacks through the network. In order to improve the effectiveness of the IDS and predictability, the CNN implements the convolution for collecting the local features, whereas a deep - layer - RNN extracts the features in this developed Hybrid DL - Based NIDS (HDL - NIDS).

In [14], a DL method was proposed. The technique was developed a two different DL algorithms. Primarily, the dataset under the DNN technique has been implemented followed by a similar dataset under CNN method. The concept of explainable Artificial Intelligence (XAI) was also employed. The authors [15] introduced three DL based misbehavior classification methods deploying CNNs and LSTM. The developed DL Classification Engines (DCLE) encompass single or multi - stage classifications achieved by DL techniques, which will be applied the vehicular edge servers. Vehicular information obtained by the RSU was pre - processed and transferred to the edge server for categorization and also 3 classification methods were developed.

In [16], the Deep Residual - CNN (DCRNN) techniques were developed to improve network security via intrusion detection that could be enhanced by the Improved Gazelle Optimization Algorithm (IGOA) system. FS should remove unrelated features existing in the network data utilized in obstacle classification methods. Vital features have been elected through the New Binary Grasshopper Optimization Algorithm (NBGOA). Sakthi and Nirmal Kumar [17] projected an effective DL method employing an optimum weight - based - DNN (OWDNN) model. Later pre - processing, the data under - sampling method must be executed by exploiting the butterfly - optimized k - means clustering (BOKMC) approach. The applicable features from the balanced dataset have been chosen through inception - V3 with multi - head attention (IV3MHA) technique. Then, the dimensionality of these chosen features will be minimized dependent upon PCA. In conclusion, the classification was accomplished by employing the OWDNN technique.

## 3. The Proposed Method

In this study, we have presented a novel SCAOCAE - IDC methodology. The developed SCAOCAE - IDC system presents a wide - ranging strategy to improve the precision and effectiveness of IDSs. The method combines diverse advanced mechanisms like Min - Max scalar normalization for data preprocessing, SCA for FS, CAE for better feature extraction and classification, and HBO for hyperparameter tuning. Fig.1 illustrates the entire flow of SCAOCAE - IDC algorithm.

### 3.1 Data Preprocessing

Initially, the SCAOCAE - IDC method applies min - max scalar normalization for data preprocessing. Min - max scaling is a normalization method commonly utilized in data preprocessing to normalize statistical features within a particular range, usually between zero and on. By rescaling data such as standardized interval, the min - max scalar confirms that each feature provides consistency to ML methods, avoiding some specific variable owing to its scale. This technique is mainly beneficial while executing with various datasets, improving the robustness and convergence of methods by mitigating the effects of differing magnitudes through diverse features, and finally providing enriched system interpretability and effectiveness.

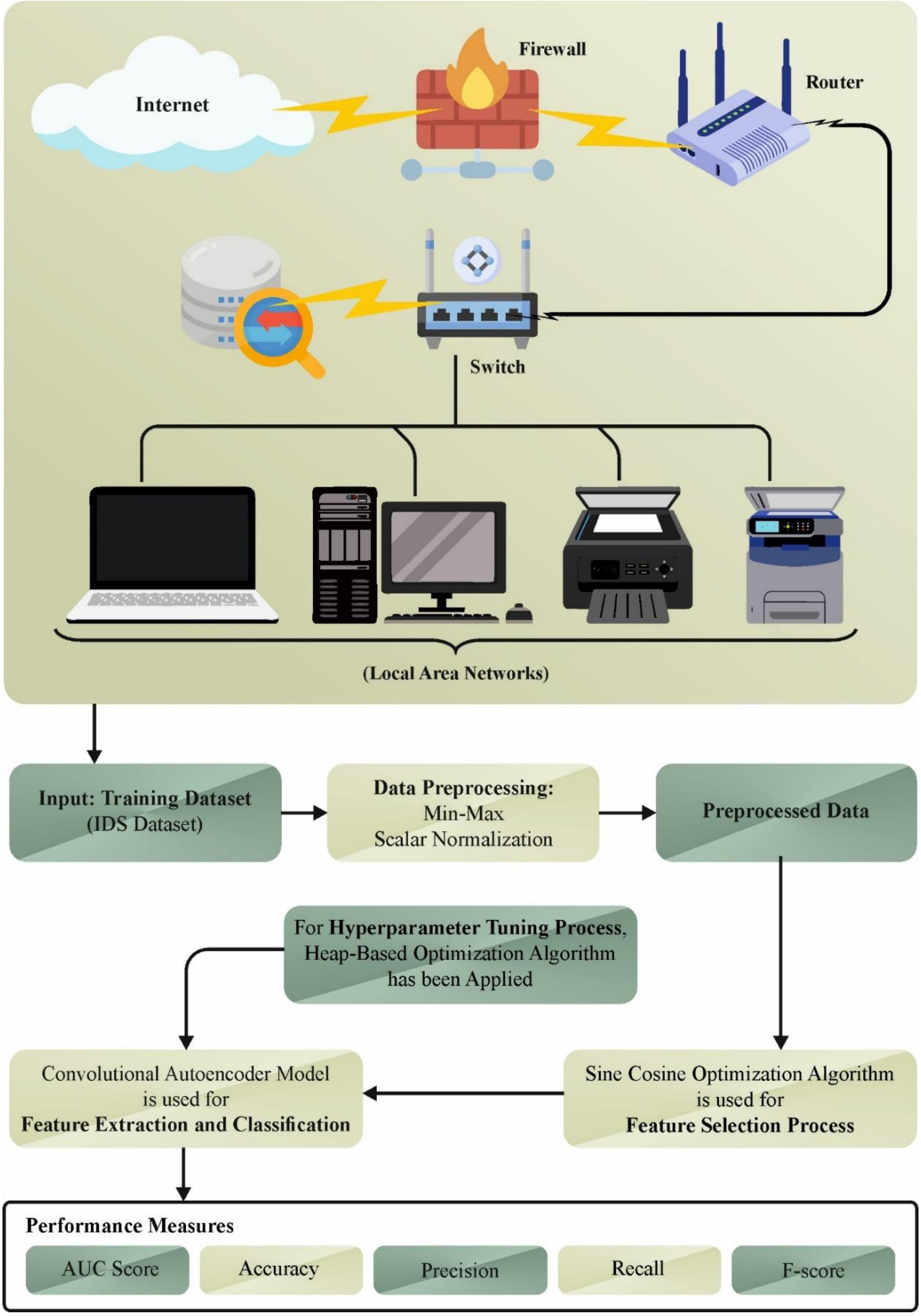

Internet
Firewall
Router
Switch
(Local Area Networks)
Input: Training Dataset (IDS Dataset)
Data Preprocessing: Min-Max Scalar Normalization
Preprocessed Data
For Hyperparameter Tuning Process, Heap-Based Optimization Algorithm has been Applied
Convolutional Autoencoder Model is used for Feature Extraction and Classification
Sine Cosine Optimization Algorithm is used for Feature Selection Process
Performance Measures
AUC Score
Accuracy
Precision
Recall
F-score

**Figure 1:** Overall flow of SCAOCAE - IDC algorithm

### 3.2 SCA based Feature Selection

The SCA can be utilized for the feature selection process. SCA is a population-based technique stimulated by the mathematical rules [18]. The main idea is to apply the behavior of sine and cosine functions for optimization method. Similar to other optimization methods, SCA begins with the initialization stage which includes a population of agents following a random manner to generate the initial solution. This agent is updated iteratively by applying the features of sine and cosine functions through random parameters.

**Updating phase**

Every solution is evolved by updating the process.

$$x_i^{(t+1)} = \begin{cases} x_i^{(t)} + A_1 \times \sin(A_2) \times |A_3 x_{best}^{(t)} - x_i^{(t)}| & r < 0.5 \\ x_i^{(t)} + A_1 \times \cos(A_2) \times |A_3 x_{best}^{(t)} - x_i^{(t)}| & r \geq 0.5 \end{cases} \quad (1)$$

In Eq. (1), $r$ represents the uniformly random integer within $(0, 1)$ and performs the transition from sine to cosine form. $x_i^{(t+1)}$ and $x_i^{(t)}$ are the $i^{th}$ $(i = 1, 2, \ldots, N)$ solution vector at $(t + 1)$ th and $t^{th}$ iteration correspondingly and $x_{best}^{(t)}$ is the best solution vector at $t^{th}$ iteration. The parameter $A_2$ is accountable for the movement of existing solutions either towards $x_{best}^{(t)}$ or outside $x_{best}^{(t)}$. The $A_3$ weight parameter could emphasize the exploration stage if $A_3 > 1$ and improve the exploitation if $A_3 < 1$.

**Balancing phase**

The balancing stage is accountable for maintaining a better balance amongst the diversification and intensification features thereby preventing the dilemma of early convergence. The parameter $A_1$ is presented by deciding the search area around the existing solution, which lies inside $x_{best}^{(t)}$ and $x_i^{(t)}$ or outside them. Consequently, $A_1$ parameter used to contribute the exploration and exploitation features in the initial and next half of the overall amount of iterations and it is given as follows:

$$A_1 = 2 - \frac{2 \times l}{L} \quad (2)$$

In Eq. (2), $l$ and $L$ are the current and maximal iterations correspondingly.

The fitness function considers the amount of attributes chosen and the classifier's accuracy. It decreases the set size of features chosen and improves the classifier accuracy. Thus, the FF is utilized to assess each solution, as follows:

$$Fitness = \alpha * ErrorRate + (1 - \alpha) * \frac{\#SF}{\#All\_F} \quad (3)$$

Where $ErrorRate$ means the classifier error values utilizing the features elected. $\alpha$ controls the importance of classifier quality and subset length and $\alpha$ is set to 0.9. $ErrorRate$ is measured as the percentage of incorrect classified to the number of classifiers made, expressed as a value between 0 and 1. ($ErrorRate$ is the complement of classifier accuracy), $\#SF$ is the number of features chosen and $\#All\_F$ is the overall quantity of attributes in the new dataset.

### 3.3 Classification using CAE Model

At this stage, the SCAOCAE - IDC method undergoes the CAE for better detection and classification processes. An AE is a feed-forward ANN whose main aim is to imitate the layers of input at the output [19]. The structure of AE is defined by the link of dual systems namely encoder and decoder. The encoding is highly liable for the input conversion from higher to a low dimension space feature, and then a decoding rebuilds the original signal from the individual code. Depending upon this outline, both networks are trained equally by modifying the weight of decoder at first and then the weight of encoder. The foremost intention of this structure is mainly dependent upon the minimization of dissimilarity among the input (original signal) and output (reconstruction).

The encoding delivers a nonlinear map of input dataset to calculate the encoded feature $e_i$:

$$e_i = \sigma(Wx_i + b), (4)$$

Whereas $b$ and $W$ signify the biases and weights, correspondingly. $\sigma$ signifies the nominated activation function. Encoding values are then decoded for reconstructing the input data $x$ of original by:

$$\hat{x}_i = \sigma(\widetilde{W}x_i + \tilde{b}), (5)$$

Here, $\tilde{b}$ and $\widetilde{W}$ defines the bias and weight value, respectively. During the stage of training, the function of error is reduced by adapting $b, \tilde{b}, W, \widetilde{W}$:

$$\arg \min_{W, b, \widetilde{W}, \tilde{b}} f\ (x, \hat{x})\ . (6)$$

Where, $f\ (\cdot)$ denotes the function of cost, which is definite as squared error or function of cross-entropy between others. With that regard, deep AE or AE is followed by multi-layered ANN which is presented for reducing the rate of error. Similarly, alterations to the classical structure of AE are planned in the work. A few instances namely CAE, Variational AE (VAE), Sparse AE (SAE), and Denoising AE (DAE).

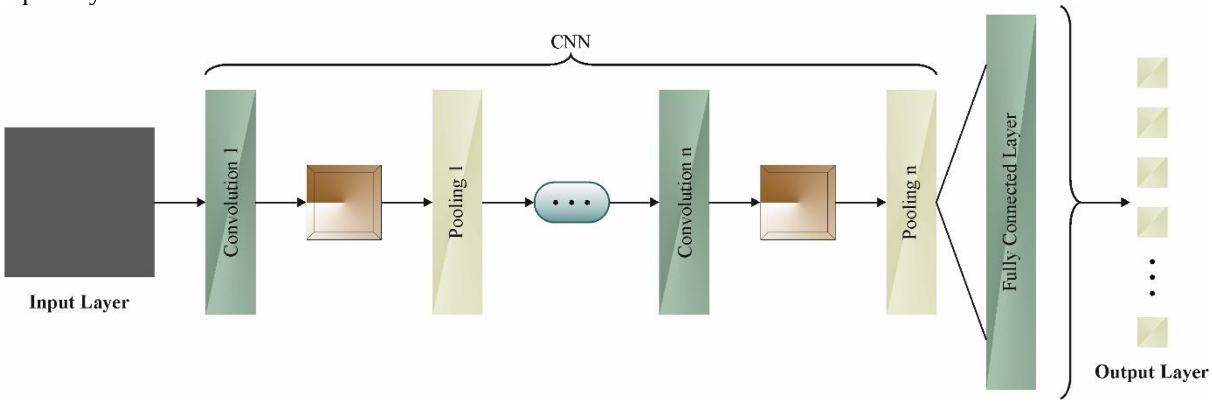


**Figure 2:** CAE architecture

The structure of CAE contains convolution and deconvolution layers in encode and decode phases, correspondingly, instead of fully connected (FC) layer. Furthermore, increasing layers are combined like pooling, and up or down sampling. The convolutional process has been implemented by descending the filter across the input signals. The subsampling or pooling layers have been situated after the convolution layer to simplify the resultant data. Dependent upon this hint, three - dimensional and two - dimensional CNN have attained effective outcomes in image and video processing in numerous fields. Fig.2 demonstrates the infrastructure of CAE.

Paper ID: SR24507124503 DOI: https://dx.doi.org/10.21275/SR24507124503

### 3.4 Hyperparameter Tuning Process

Finally, HBO fine - tunes hyperparameters for improved system performance. HBO algorithm is inspired by the social attitudes of humans towards organizational hierarchy [20]. This technique is emulated by the corporate rank hierarchy (CRH). The presented method is based on the hierarchical arrangement of search candidates on fitness using the corporate rank hierarchy concept. Using the heap-based data structure, the hierarchy is constructed. In addition to the modelling of corporate rank hierarchy, the overall concept consists of three different stages: (i) modelling the cooperation amongst the direct manager and the subordinators; (ii) modelling the employee interaction, and (iii) modelling the self-contribution of subordinator to achieve the desired function. The above steps are briefly discussed below.

Using nonlinear tree-shaped data structure, corporate rank hierarchy is constructed. In the suggested method, the enhanced CRH is regarded as a swarm. The heap node represents the searching agent within the search range, and the fitness function indicates the master key to the heap nodes. By taking the heap node, the population index of search candidate can be defined.

**Modeling with Direct Manager Interaction**

Senior leadership uniformly enforces laws and regulations on employees in big organizations with central organizational structure, and employees should follow directions from their managers. The next phase is mathematically described by changing the locations of search candidates:

$$x_i^k(t+1) = B^k + \gamma(2r-1)\left|B^k - x_i^k(t)\right| \quad (7)$$

In Eq. (7), $x$ indicates the search agent's location, B indicates the parental node, and the existing iteration and the component vectors are represented by t and k, correspondingly.

$$\lambda^k = 2r - 1 \quad (8)$$

Where the term $(2r-1)$ represents the $k^{th}$components of $\gamma$ vector, and is arbitrarily generated, $r$ denotes uniformly distributed random parameter lies between zero and one.

$$\gamma = \left|2 - \frac{(tmod_C^T)}{\frac{T}{4c}}\right| \quad (9)$$

In Eq. (9), the maximal iteration counts is $T$, and $C$ indicates the adjustable parameter:

$$C = \frac{T}{25} \quad (10)$$

**Modeling of Interaction between the Subordinators**

In a certain organization, colleagues who work together as subordinator accomplishes official responsibility. At the same position, the nodes in the heap are considered as colleagues.

$$x_i^k(t+1) = \begin{cases} S_r^k + \gamma\lambda^k\left|S^k + x_i^k(t)\right|, f(S_r) < f\left(x_i^k(t)\right) \\ x_i^k(t) + \gamma\lambda^k\left|S^k + x_i^k(t)\right|, f(S_r) \geq f\left(x_i^k(t)\right) \end{cases} \quad (11)$$

Eq. (11) represents the updating location $(x)$ of search agent based on the random teammate selection $(S_r)$.

**Modeling Self-Contribution of Employees**

Eq. (12) defines how the employees in the organization are self-contributing:

$$x_i^k(t+1) = x_{i(t)} \quad (12)$$

**Updating Position**

A roulette wheel mechanism is used for balancing the exploitation and exploration phases. By applying the $P_1, P_2$, and $P_3$ probabilities, the equilibrium between both stages is established. Using the first probability $P_1$, the location of search agent in the population was updated:

$$P_1 = 1 - \frac{t}{T} \quad (13)$$

The second proportion $P_2$, is defined as:

$$P_2 = P_1 + \frac{1-P_1}{2} \quad (14)$$

Using Eq. (15), the probability $P_3$ was defined:

$$P_3 = P_2 + \frac{1-P_1}{2} = 1 \quad (15)$$

The upgrading position equation for the HBO is presented in Eq. (16).

$$x_i^k\,(t+1) = \begin{cases} x_i^k(t), P < P_1 \\ B^k + \gamma\lambda^k\left|B^k - x_i^k(t)\right|,\ P_2 < P < P_3 \\ S_r^k + \gamma\lambda^k|S_r^k + x^k\ (t)\ |, P_2 < P < P_3\ and\ f\ (S_r)\ \geq f\left(x_i^k(t)\right) \\ x_i^k + \gamma\lambda^k\left|S_r^k - x_i^k(t)\right|, P_2 < P < P_3\ and\ f\ (S_r)\ < f\left(x_i^k(t)\right) \end{cases} \quad (16)$$

Where $P$ indicates any number lies within the range $[0, 1]$.

The HBO method derives an FF to obtain superior classifier accuracy. It defines a positive integer to characterize the high efficiency of candidate solution. Here, the decline of the classifier error rate is supposed as FF.

$$fitness(x_i) = ClassifierErrorRate(x_i) = \frac{No.of\ misclassified\ samples}{Total\ No.of\ samples} * 100 \quad (17)$$

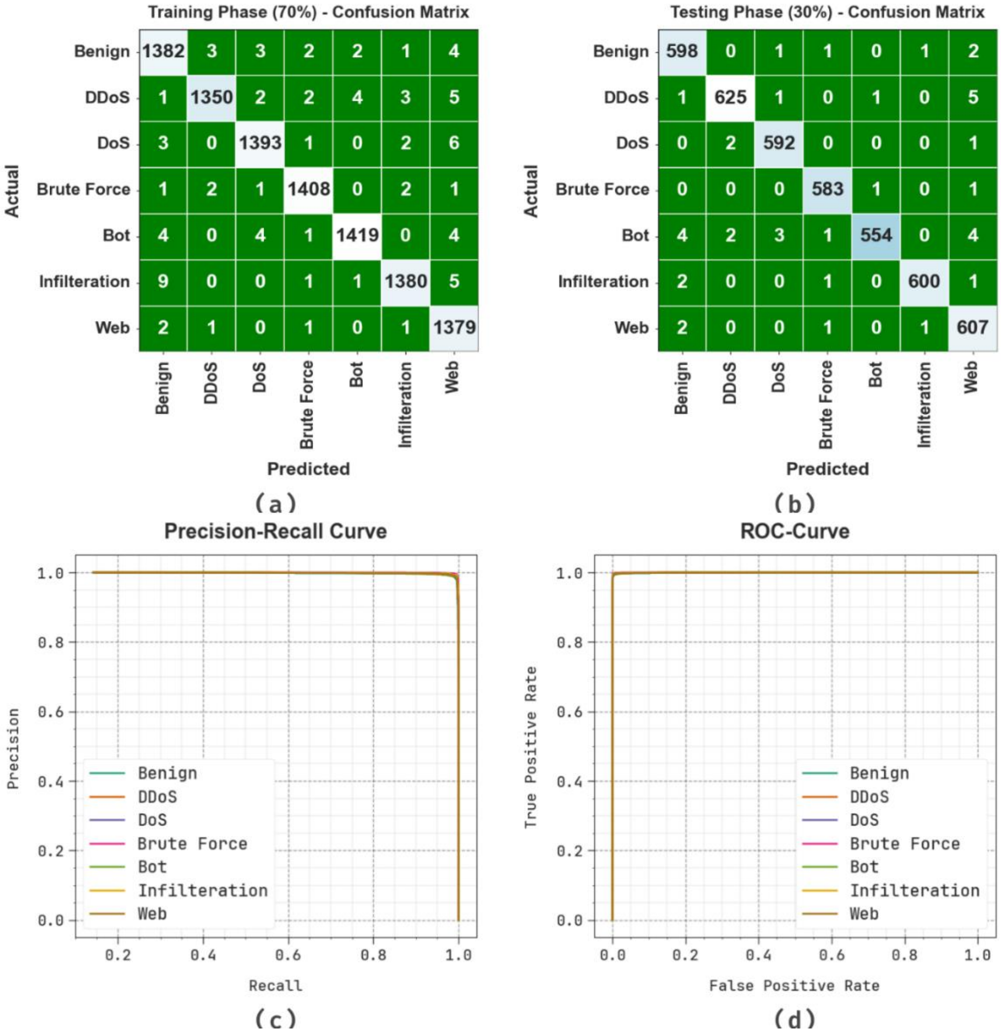
## 4. Result Analysis and Discussion

The performance study of the SCAOCAE - IDC algorithm undergoes utilizing the CICIDS2018 dataset [21]. The SCAOCAE - IDC technique uses CGAN to balance the dataset. It has 13995 samples with 7 classes as illustrated in Table 1. The SCAOCAE - IDC system has chosen 46 features from the available 80 features.

**Table 1:** Details on database

| Classes | No. of Instances |
|---|---|
| Benign | 2000 |
| DDoS | 2000 |
| DoS | 2000 |
| Brute Force | 2000 |
| Bot | 2000 |
| Infilteration | 2000 |
| Web | 1995 |
| Total Instances | 13995 |

Fig.3 exposes the classifier outcome of the SCAOCAE - IDC algorithm on test database. Figs.3a - 3b represents the confusion matrices provided by the SCAOCAE - IDC system on 70: 30 of TRAS/TESS. The experimental value implied that the SCAOCAE - IDC technique appropriately recognizes the intrusions under all classes. In addition, Fig.3c illustrates the PR value of the SCAOCAE - IDC model. The experimental value reported that the SCAOCAE - IDC model has reaches greater PR values under 7 classes. However, Fig.3d illustrates the ROC curve of the SCAOCAE - IDC algorithm. The experimental value depicted that the SCAOCAE - IDC algorithm has resulted in adept solutions with higher values of ROC under 7 classes.


Training Phase (70%) - Confusion Matrix

| Actual \ Predicted | Benign | DDoS | DoS | Brute Force | Bot | Infilteration | Web |
|---|---|---|---|---|---|---|---|
| Benign | 1382 | 3 | 3 | 2 | 2 | 1 | 4 |
| DDoS | 1 | 1350 | 2 | 2 | 4 | 3 | 5 |
| DoS | 3 | 0 | 1393 | 1 | 0 | 2 | 6 |
| Brute Force | 1 | 2 | 1 | 1408 | 0 | 2 | 1 |
| Bot | 4 | 0 | 4 | 1 | 1419 | 0 | 4 |
| Infilteration | 9 | 0 | 0 | 1 | 1 | 1380 | 5 |
| Web | 2 | 1 | 0 | 1 | 0 | 1 | 1379 |

Testing Phase (30%) - Confusion Matrix

| Actual \ Predicted | Benign | DDoS | DoS | Brute Force | Bot | Infilteration | Web |
|---|---|---|---|---|---|---|---|
| Benign | 598 | 0 | 1 | 1 | 0 | 1 | 2 |
| DDoS | 1 | 625 | 1 | 0 | 1 | 0 | 5 |
| DoS | 0 | 2 | 592 | 0 | 0 | 0 | 1 |
| Brute Force | 0 | 0 | 0 | 583 | 1 | 0 | 1 |
| Bot | 4 | 2 | 3 | 1 | 554 | 0 | 4 |
| Infilteration | 2 | 0 | 0 | 1 | 0 | 600 | 1 |
| Web | 2 | 0 | 0 | 1 | 0 | 1 | 607 |

**Figure 3:** Classifier outcomes of (a - b) Confusion matrices and (c - d) PR and ROC curves

The intrusion recognition results of the SCAOCAE - IDC methodology are investigated briefly in Table 2 and Fig.4. The results outlined that the SCAOCAE - IDC technique appropriately recognizes the intrusions under all classes. With 70%TRAS, the SCAOCAE - IDC technique provides average $accu_y$ of 99.75%, $prec_n$ of 99.13%, $reca_l$ of 99.13%, $F_{score}$ of 99.13%, and $AUC_{score}$ of 99.49%. Additionally, with 30%TESS, the SCAOCAE - IDC approach gains average $accu_y$ of 99.75%, $prec_n$ of 99.13%, $reca_l$ of 99.13%, $F_{score}$ of 99.13%, and $AUC_{score}$ of 99.49%.

**Table 2:** Intrusion recognition outcome of SCAOCAE - IDC technique under 70: 30 of TRAS/TESS

| **Classes** | $Accu_y$ | $Prec_n$ | $Reca_l$ | $F_{Score}$ | $AUC_{Score}$ |
|---|---|---|---|---|---|
| **TRAS (70%)** | | | | | |
| Benign | 99.64 | 98.57 | 98.93 | 98.75 | 99.34 |
| DDoS | 99.77 | 99.56 | 98.76 | 99.16 | 99.34 |
| DoS | 99.78 | 99.29 | 99.15 | 99.22 | 99.51 |
| Brute Force | 99.85 | 99.44 | 99.51 | 99.47 | 99.70 |
| Bot | 99.80 | 99.51 | 99.09 | 99.30 | 99.50 |
| Infilteration | 99.74 | 99.35 | 98.85 | 99.10 | 99.37 |
| Web | 99.69 | 98.22 | 99.64 | 98.92 | 99.67 |
| **Average** | **99.75** | **99.13** | **99.13** | **99.13** | **99.49** |
| **TESS (30%)** | | | | | |
| Benign | 99.67 | 98.52 | 99.17 | 98.84 | 99.46 |
| DDoS | 99.71 | 99.36 | 98.74 | 99.05 | 99.31 |
| DoS | 99.81 | 99.16 | 99.50 | 99.33 | 99.68 |
| Brute Force | 99.86 | 99.32 | 99.66 | 99.49 | 99.77 |
| Bot | 99.62 | 99.64 | 97.54 | 98.58 | 98.74 |
| Infilteration | 99.86 | 99.67 | 99.34 | 99.50 | 99.64 |
| Web | 99.57 | 97.75 | 99.35 | 98.54 | 99.48 |
| **Average** | **99.73** | **99.06** | **99.04** | **99.05** | **99.44** |

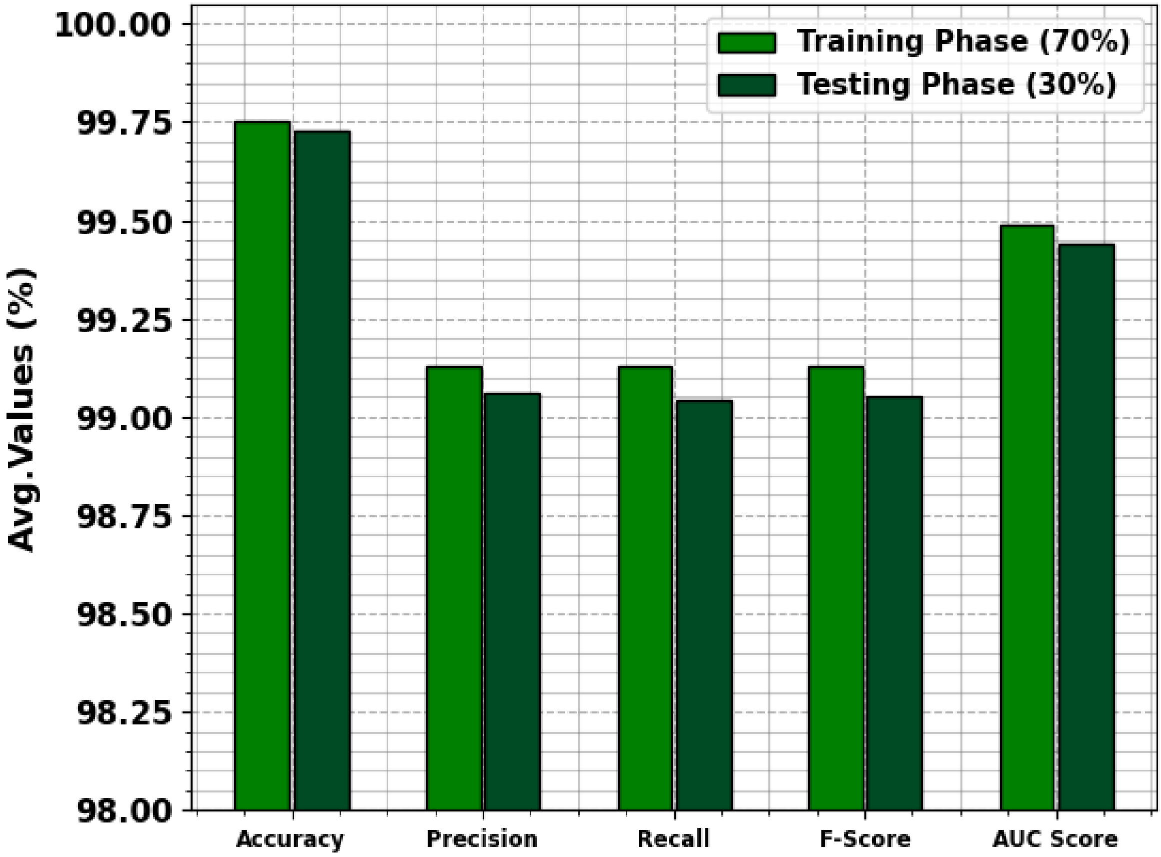


**Figure 4:** Average of SCAOCAE - IDC technique under 70: 30 of TRAS/TESS

The performance of the SCAOCAE - IDC approach is graphically presented in Fig.5 in the form of training accuracy (TRAA) and validation accuracy (VALA) curves. The outcome exhibits a useful interpretation of the behaviour of the SCAOCAE - IDC approach over several epoch counts, demonstrating its learning process and generalization abilities. Remarkably, the figure infers a steady enhancement in the TRAA and VALA with progress in epochs. It ensures the adaptive nature of the SCAOCAE - IDC model in the pattern recognition process on both TRA and TES data. The rising trend in VALA summarizes the ability of the SCAOCAE - IDC model to adapt to the TRA data and also excels in offering accurate classification of unseen data, pointing out robust generalized capabilities.

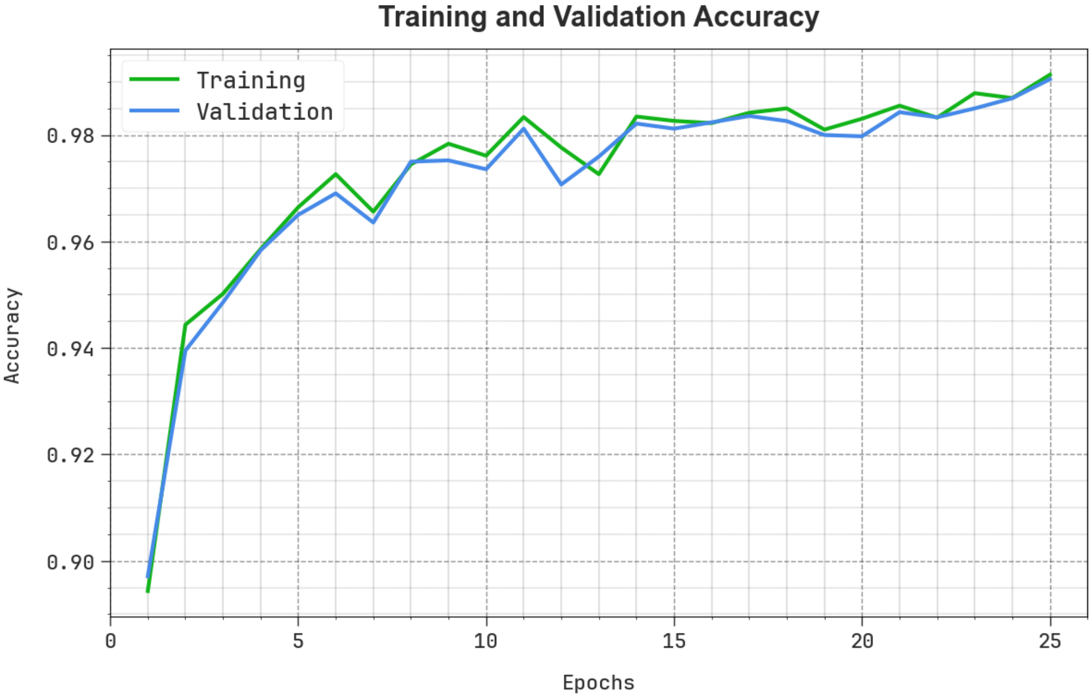


**Figure 5:** $Accu_y$ curve of the SCAOCAE - IDC technique

Fig.6 exposes a complete representation of the training loss (TRLA) and validation loss (VALL) outcomes of the SCAOCAE - IDC system under distinct epochs. The progressive reduction in TRLA highlights the SCAOCAE - IDC system optimizing the weights and decreasing the classification error on the TRA and TES data. The figure inferred a clear understanding of the SCAOCAE - IDC method's association with the TRA data, highlighting its proficiency in capturing patterns from both datasets. Remarkably, the SCAOCAE - IDC model continually greats its parameters in minimizing the differences among the predictive and real TRA class labels.

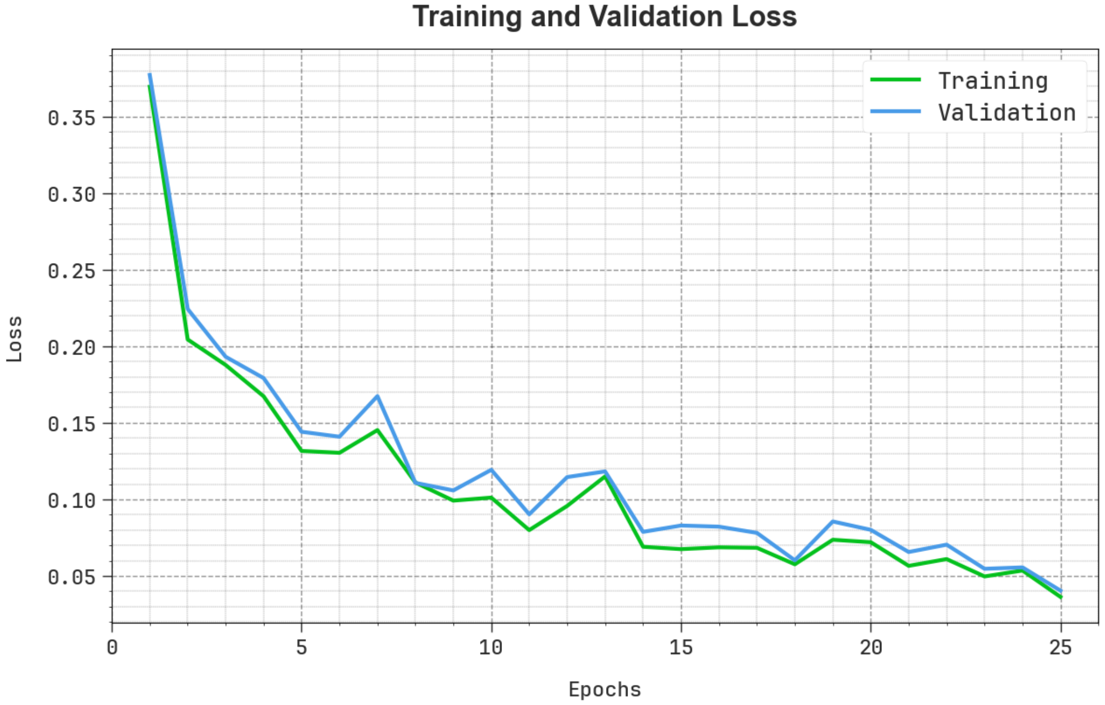


**Figure 6:** Loss curve of the SCAOCAE - IDC technique

In Table 3 and Fig.7, the results of the SCAOCAE - IDC technique are compared with existing models [22]. The results indicate that the SCAOCAE - IDC method gains enhanced performance over other approaches. Based on $accu_y$, the SCAOCAE - IDC algorithm offers higher $accu_y$ of 99.75% while the DT, RF, XT, AdaBoost, LGBM, and XGB models obtain lower $accu_y$ of 98.70%, 98.40%, 98.30%, 97.80%, 98.80%, and 98.90%, correspondingly. Also, based on $prec_n$, the SCAOCAE - IDC approach provides superior $prec_n$ of 99.13% while the DT, RF, XT, AdaBoost, LGBM, and XGB systems attain lesser $prec_n$ of 96.67%, 97.43%, 97.35%, 97.74%, 97.83%, and 96.97%, correspondingly.

**Table 3:** Comparative outcome of SCAOCAE - IDC system with recent other models

| Model | $Accu_y$ | $Prec_n$ | $Reca_l$ | $F_{score}$ |
|---|---|---|---|---|
| SCAOCAE - IDC | 99.75 | 99.13 | 99.13 | 99.13 |
| Decision tree | 98.70 | 96.67 | 97.67 | 97.67 |
| Random forest | 98.40 | 97.43 | 95.43 | 96.43 |
| Extra tree | 98.30 | 97.35 | 97.35 | 98.05 |
| AdaBoost | 97.80 | 97.74 | 97.65 | 97.80 |
| LGBM | 98.80 | 97.83 | 96.83 | 97.83 |
| XGB | 98.90 | 96.97 | 96.98 | 96.97 |

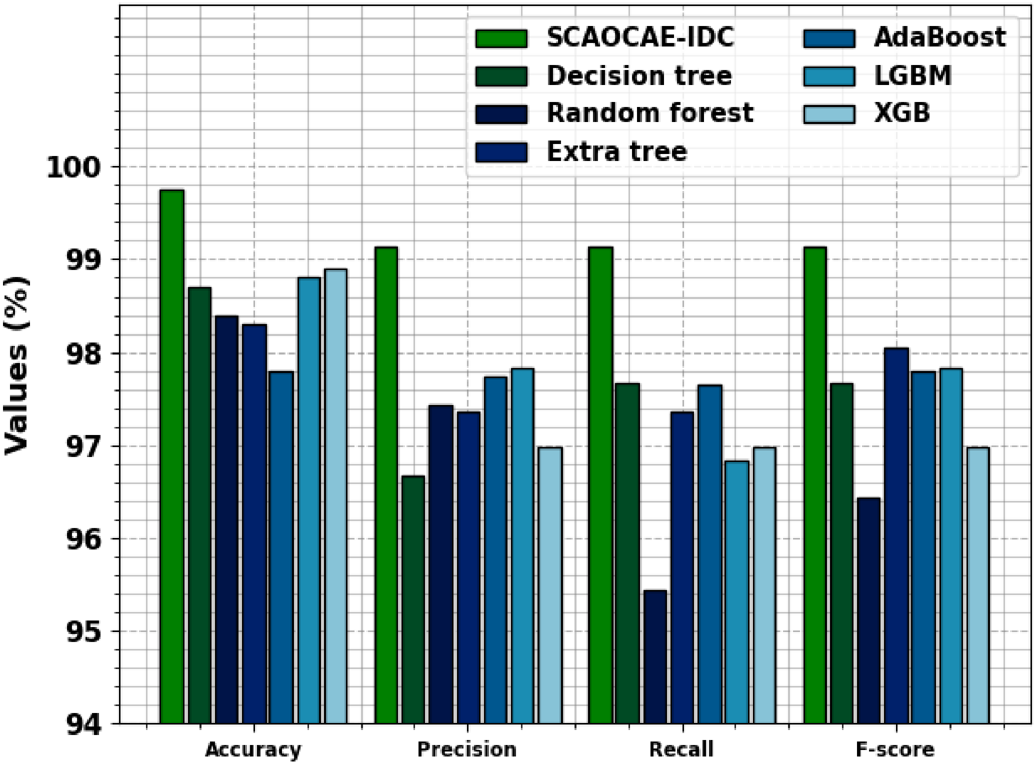


**Figure 7:** Comparative analysis of SCAOCAE - IDC technique with other approaches

Finally, based on $F_{score}$, the SCAOCAE - IDC methodology provides maximal $F_{score}$ of 99.13% while the DT, RF, XT, AdaBoost, LGBM, and XGB systems reached minimal $F_{score}$ of 97.67%, 96.43%, 98.05%, 97.80%, 97.83%, and 96.97%, correspondingly. Therefore, the SCAOCAE - IDC technique can be applied for improved recognition of the intrusions.

## Conclusion

In this study, we have presented a novel SCAOCAE - IDC methodology. The developed SCAOCAE - IDC system presents a wide - ranging strategy to improve the precision and effectiveness of IDSs. The method combines diverse advanced mechanisms like Min - Max scalar normalization for data preprocessing, SCA for FS, CAE for better feature extraction and classification, and HBO for hyperparameter tuning. The Min - Max scalar makes sure of robust data normalization, SCA increasingly chooses main features, CAE capably captures complex patterns in the data, and HBO fine - tunes hyperparameters for improved system performance. By employing the synergistic combination of such modules, the presented SCAOCAE - IDC algorithm indicates considerable outcomes for increasing the reliability and accuracy of IDSs and classification systems.

## References

[1] Lin, Y. D., Liu, Z. Q., Hwang, R. H., Nguyen, V. L., Lin, P. C. and Lai, Y. C., 2022. Machine learning with variational AutoEncoder for imbalanced datasets in intrusion detection. *IEEE Access*, *10*, pp.15247 - 15260.

[2] Panigrahi, R., Borah, S., Bhoi, A. K., Ijaz, M. F., Pramanik, M., Kumar, Y. and Jhaveri, R. H., 2021. A consolidated decision tree - based intrusion detection system for binary and multiclass imbalanced datasets. *Mathematics*, *9* (7), p.751.

[3] Seo, J. H. and Kim, Y. H., 2018. Machine - learning approach to optimize smote ratio in class imbalance dataset for intrusion detection. *Computational intelligence and neuroscience*, *2018*.

[4] Rao, Y. N. and Suresh Babu, K., 2023. An Imbalanced Generative Adversarial Network - Based Approach for Network Intrusion Detection in an Imbalanced Dataset. *Sensors*, *23* (1), p.550.

[5] Rani, M. and Gagandeep, 2022. Effective network intrusion detection by addressing class imbalance with deep neural networks multimedia tools and applications. *Multimedia Tools and Applications*, *81* (6), pp.8499 - 8518.

[6] Pimsarn, C., Boongoen, T., Iam - On, N., Naik, N. and Yang, L., 2022. Strengthening intrusion detection system for adversarial attacks: improved handling of imbalance classification problem. *Complex & Intelligent Systems*, *8* (6), pp.4863 - 4880.

[7] Tran, N., Chen, H., Jiang, J., Bhuyan, J. and Ding, J., 2021. Effect of Class Imbalance on the Performance of Machine Learning - based Network Intrusion Detection. *International Journal of Performability Engineering*, *17* (9).

[8] Amalapuram, S. K., Reddy, T. T., Channappayya, S. S. and Tamma, B. R., 2021, October. On handling class imbalance in continual learning based network intrusion detection systems. In *The First International Conference on AI - ML - Systems* (pp.1 - 7).

[9] Seo, J. H. and Kim, Y. H., 2018. Machine - learning approach to optimize smote ratio in class imbalance dataset for intrusion detection. *Computational intelligence and neuroscience*, *2018*.

[10] Zhang, H., Huang, L., Wu, C. Q. and Li, Z., 2020. An effective convolutional neural network based on SMOTE and Gaussian mixture model for intrusion detection in imbalanced dataset. *Computer Networks*, *177*, p.107315.

[11] Ramu, C. K., Rao, T. S. and Rao, E. U. S., 2024. Attack classification in network intrusion detection system based on optimization strategy and deep learning methodology. *Multimedia Tools and Applications*, pp.1 - 23.

[12] Umer, M., Sadiq, S., Karamti, H., Alhebshi, R. M., Alnowaiser, K., Eshmawi, A. A., Song, H. and Ashraf, I., 2022. Deep Learning - Based Intrusion Detection Methods in Cyber - Physical Systems: Challenges and Future Trends. *Electronics*, *11* (20), p.3326.

[13] Qazi, E. U. H., Faheem, M. H. and Zia, T., 2023. HDLNIDS: Hybrid Deep - Learning - Based Network Intrusion Detection System. *Applied Sciences*, *13* (8), p.4921.

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

[15] Alladi, T., Kohli, V., Chamola, V. and Yu, F. R., 2023. A deep learning based misbehavior classification scheme for intrusion detection in cooperative intelligent transportation systems. *Digital Communications and Networks*, *9* (5), pp.1113 - 1122.

[16] Kumar, G. S. C., Kumar, R. K., Kumar, K. P. V., Sai, N. R. and Brahmaiah, M., 2024. Deep residual convolutional neural Network: An efficient technique for intrusion detection system. *Expert Systems with Applications*, *238*, p.121912.

[17] Sakthi, K. and Nirmal Kumar, P., 2023. A novel attention - based feature learning and optimal deep learning approach for network intrusion detection. *Journal of Intelligent & Fuzzy Systems*, *45* (3), pp.5123 - 5140.

[18] Rizk - Allah, R. M. and Hassanien, A. E., 2023. A comprehensive survey on the sine–cosine optimization algorithm. *Artificial Intelligence Review*, *56* (6), pp.4801 - 4858.

[19] Rodriguez, M. A., Sotomonte, J. F., Cifuentes, J. and Bueno - López, M., 2020, September. Power quality disturbance classification via deep convolutional auto - encoders and stacked LSTM recurrent neural networks. In *2020 International Conference on Smart Energy Systems and Technologies (SEST)* (pp.1 - 6). IEEE.

[20] Gör, H., 2024. Feasibility of Six Metaheuristic Solutions for Estimating Induction Motor Reactance. *Mathematics*, *12* (3), p.483.

[21] https: //registry. opendata. aws/cse - cic - ids2018/

[22] Ogobuchi Okey, D.; Sarah Maidin, S.; Adasme, P.; Lopes Rosa, R.; Saadi, M.; Carrillo Melgarejo, D.; Zegarra Rodríguez, D. BoostedEnML: Efficient Technique for Detecting Cyberattacks in IoT Systems Using Boosted Ensemble Machine Learning. Sensors 2022, 22, 7409. https: // doi. org/10.3390/s22197409