International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Intrusion Detection by Stacked Deep Ensemble Model with Entropy and Correlation Feature Set

Sravanthi Godala[1], Dr. M. Sunil Kumar[2]

Submitted: 09/11/2022 **Accepted:** 05/02/2023

Abstract*:* Nowadays, in order to improve the routine activities, interoperability and interconnectivity of computing systems are extensively used. In addition, it creates a way to vulnerabilities that are far beyond the reach of human control. Due to the vulnerabilities, data transfer must include cyber-security measures. Secure connectivity demands improvements to security mechanisms to counter emerging security risks and security systems to mitigate the threats. This paper proposes Intrusion Detection by Stacked Deep Ensemble Model (IDSDEM) which has three working stages. Initially, in the pre-processing stage, data normalization process is conducted to reduce the data redundancy and increases the consistency of data for further process. Afterwards feature extraction stage takes place where the features such as entropy based as well as improved correlation based features were extracted. Finally, intrusion detection is conducted where a stacked deep ensemble model which includes the classifiers like Deep Belief Network (DBN), Deep Maxout Network and Customized Convolutional Neural Network (CCNN) is employed to provide effective intrusion detection. The outcomes demonstrated that the developed IDSDEM can offer superior performance with respect to detection accuracy, precision and other measures.

Keywords: *Intrusion Detection, Stacked Deep Ensemble Model, Data normalization, Improved correlation based feature, Deep Belief Network (DBN), Deep Maxout Network, Customized Convolutional Neural Network (CCNN)***.**

1. Introduction

The need for local environmental data has grown dramatically along with the Internet of Things and real-time big data's quick development. The need for Wireless sensor Network (WSN) devices with affordable node costs and simple setup will progressively increases. WSN products are capable of evading typical detection techniques [9][10]. They significantly decreased the cost of ecological assessment and the labor-intensive procedure of conventional testing approaches. Since it was created, the WSN has been investigated extensively by researchers and utilized by industry. Environmental monitoring, military activities, including data positioning are examples of the typical applications [11][12][13]. The openness of WSNs implementation areas and wireless communication's capacity to broadcast, make the network susceptible to external threats or intrusions, greatly growing the exposure to threats that harm the accessibility of data systems in the network management system. Since it involves gathering and analyzing network data to find unusual activity on the network, intrusion detection systems are frequently used to safeguard the networks from security threats [14][15][16].

WSNs intrusion detection generally divided into two groups by applying various detection techniques: misuse detection as well as anomaly detection. The first is a mathematical model-based detection approach in which a standardized network model is created with a typical network behavior pattern and it's evaluated whether certain network behavior feature values vary from average values [17][18][19]. If the limiting threshold has been surpassed, it is considered that there has been an intrusion. Anomaly detection systems encompass data mining-based anomaly detection, machine learning-based anomaly detection, as well as clustering-based anomaly detection; the latter is an information-based intrusion detection method that creates a data state information source for known threats network behaviour and creates one or many matching patterns for every intrusion. If the matching patterns were located in the data base and correlate with user behaviour, the extant intrusion patterns could be immediately recognized [20][21][22].

Because of the benefits of self-learning, categorization, and adaptability, neural networks (NNs) has drawn a significant range of researchers to explore the intrusion detection method depending on NNs and have produced very good results. Lately, a deep learning (DL) technique for network intrusion detection was suggested, which

[1] *Research Scholar, Department of CSE, JNTUA, Ananthapuramu 515002, AP, India Email:vmsravanthi@gmail.com*
[2] *Professor, Department of CSE, Sree Vidyanikethan Engineering College (Autonomous), Tirupati 517102,AP, India Email:sunilmalchi1@gmail.com*

minimizes sample training time while maintaining good precision as well as detection rate [23][24]. An intrusion detection technique built on recurrent neural networks (RNNs) also examined the effectiveness in binary classification as well as multi-classification, along with the influence of a varied count of neurons & learning rates on system performance. AI has several benefits and also is extensively employed in every sector that utilises massive data sets. AI was utilized in many applications and it is not restricted to modelling predictions, detecting abnormalities, identifying attacks, malware, and frauds, and so on [25][26]. For that reason we have also developed novel IDS which include the following contributions.

- ✓ Prioposing an effective intrusion detection system named IDSDEM via introducing a Customized Convolutional Neural Network (CCNN) along with the conventional Deep Belief Network (DBN) and Deep Maxout Network.
- ✓ To provide appropriate detection model via determining the feature set with improved correlation-based features along with the entropy-based features.

Organization of this paper is given as follows: Few recently published literatures related to IDS is reviewed in section 2, proposed IDSDEM's detailed methodology is provided in section 3, Implementation findings is described in section 4, conclusions of this paper is given in section 5, Afterwards this research work's few references were given.

2. Literature Survey

Few recently published literatures related to IDS are reviewed below.

In 2022, Gaoyuan Liu et al [1] suggested WSN intelligent IDS using the k-Nearest Neighbor technique (kNN) in machine learning (ML) as well as the arithmetic optimization algorithm (AOA) in evolutionary computation to create an edge intelligence structure that particularly undertakes intrusion detection when the WSN meets a DoS attack. To improve the model's accuracy, a parallel approach was utilized to improve communication across populations and also the Lévy flight approach was employed to adjusting the optimization. The suggested PL-AOA method works effectively in the benchmark functionality test but also successfully assures the development of the kNN classifier.

In 2021, Maheswari et al [2] creates a new safe unequal clustering method with intrusion detection (SUCID) to accomplish QoS metrics such as energy, life span, as well as security.At beginning, the suggested model employs an adaptive neural fuzzy dependent clustering approach to choose tentative cluster heads (TCHs) according to 3 input parameters: residual energy, distance to neighbours and distance to base station (BS). The TCHs subsequently compete for the last CHs, and also the best CHs are chosen utilizing the deer hunting optimization (DHO) method. Utilizing residual energy, node degree, distance to BS, node centralization, and connection quality, the DHO-dependent clustering algorithm develops a fitness function. The cluster management phase was used for load balancing to increase the effectiveness of the suggested technique even further. Lastly, in order to establish safety in a cluster-dependent WSN, a successful IDS depending on a DBN was run on the CHs to recognize the existence of attackers in the network.

In 2021, Gowdhaman et al [3] proposed a deep neural network-based intrusion detection system (DNN). To discover intrusions, the cross-correlation operation was utilised to choose the best features within the dataset, and also the chosen parameters have been used as utilised as building blocks for DNN structures. The research findings showed that the suggested DNN outperforms established ML models like support vector machine (SVM), random forest as well as decision tree in identifying assaults.

In 2020, Shaimaa Ahmed Elsaid et al [4] presented an optimized collaborative ID termed OCIDS for WSNs. It employs an updated artificial bee colony optimization technique to optimise the hierarchical IDS implemented to WSNs in terms of both resource usage as well as intrusion detection accuracy. Furthermore, the suggested system enhances the weighted SVM technique to increase identification accuracy but also decrease false alarm rates. Since every sensor node, CH, as well as BS in hierarchical WSNs has a unique viewpoint on the network, interaction between them is taken into account in the suggested OCIDS approach to enable highly accurate intrusion detection.

In 2020,Hitesh Mohapatra et al [5] suggested an Man In The Middle - IDS (MITM-IDS) paradigm for detecting attacks, isolation, as well as reconfiguration for affected nodes. The IDS approach aids in training the nodes for potential threats. The experiment reveals a productivity rating of 89.147% while performing MITM attacks. This research aims to establish an attack-tolerant IDS.

In 2020,Mukaram Safaldin et al [6] presented an improved IDS depending on the improved binary grey wolf optimizer and SVM (GWOSVM-IDS). The GWOSVM-IDS experimented with three, five, and seven wolves to determine the optimal count of wolves. The suggested technique attempts to improve intrusion detection accuracy as well as detection rate while decreasing processing time in the WSN context by lowering false alarm rates as well as the amount of features produced by IDSs. According to the findings, the suggested GWOSVM-IDS having seven wolves surpasses the other suggested as well as related algorithms.

In 2020,Wenjie Zhang et al [7] suggested a hierarchical IDS which groups nodes in a WSN as per their functions. Furthermore, to enhance the detection rate of WSN IDS's abnormal behavior and lower the false alarm rate, the utilization of the classification method of kernel extreme learning machine (KELM), just after the Mercer Property to synthesize multi-kernel functions was examined in this work. Through testing as well as implementing the multi-kernel function but also building a MKELM to WSN IDS, the best linear combination was accomplished. The experimental outcomes reveal that the approach not only ensures high detection accuracy but also significantly decreases detection time, making it ideally adapted for resource-constrained WSNs.

In 2020, Shashank Gavel et al [8] developed a method that combines multi-variable kernel density assessment with decentralised computation. This integration examines the individual likelihood of data availability and computes the Probability Density Function's (PDF's) global value. Pearson's divergence (PE) was used for effective in-network ID and assessment with a minimal False Positive Rate (FPRs). The estimate of PE divergence was performed utilizing various distributed computing approaches. In order to ensure effective functionality, the value of PDFs was estimated over time. Additionally an entropy-dependent strategy based on centralised computing was proposed.

Table 1: Review on few recent publications related to IDS

Citation	Methods	Advantages	Drawbacks
Gaoyuan Liu et al [1]	PL-AOA	• High accuracy • Has good practical application significance.	• Time complexity was higher than kNN
Maheswari et al [2]	SUCID	• Obtained maximum energy efficiency • Provide maximum lifetime	• Performance still can be enhanced via hyper parameter tuning process
Gowdhaman et al [3]	DNN	• Handle imbalanced attacks. • Achieve better performance	• Accuracy can be affected by large count of attacks
Shaimaa Ahmed Elsaid et al [4]	OCIDS	• Highest detection rate • Lower computational complexity	• An adaptive feature selection scheme should be used to improve decision accuracy
Hitesh Mohapatra et al [5]	MITM-IDS	• Less complexity • Take less time to detect malicious behaviour	• Other factors like energy consumption, detection rate etc was not considered.
Mukaram Safaldin et al [6]	GWOSVM-IDS	• Lower execution time • Effectively predict unknown classes	• Performance still can be enhanced via using other classifiers
Wenjie Zhang et al [7]	MK-ELM	• Reduce the energy usage of nodes • Less false alarm rate	• Detection of many intrusion patterns should be increased
Shashank Gavel et al [8]	Pearson’s divergence	• Achieves promising performance • Highly robust	• Detecting intrusion in high-dimensional data is difficult

According to the review, conventional models lag in performance due to programming challenges. Furthermore, performance suffers since it is unable to address dynamic challenges, while ML models function well in intrusion detection. However, performances in regards of detection accuracy as well as false alarm reduction may be enhanced. Although numerous Machine ML strategies are employed in IDS, their performance when dealing with imbalanced attacks was inadequate. For that reason we introduced a novel IDS named IDSDEM which is described below.

3. Methodology

WSNs are composed of a significant count of sensor nodes that collect and send information to a central location. However, because of resource restricted nodes, installation tactics, as well as communication channels, WSNs face various security concerns. As a result, detecting unauthorized access is crucial for enhancing WSN security. Network IDSs give such features to the network, which is unavoidable for every communication network. As a consequence a novel IDS is developed in this work that has three working stages like pre-processing, feature extraction as well as intrusion detection. Initially, data normalization process is conducted in the pre-processing stage which reduces the data redundancy and increases the consistency of data for further process. Afterwards entropy based features and improved correlation based features were extracted from the pre-processed data. Finally the intrusion detection will take place by stacked deep learning model which combines the models like Deep Maxout, Deep Belief Network (DBN) and Customized Convolutional Neural Network (CCNN). Firstly extracted features were provided to the classifiers like Deep Maxout and DBN as input. In order to provide higher detection accuracy, the outcomes from the 2 classifiers were given to CCNN as input. Finally the detection or classification outcomes were provided from CCNN. Proposed IDSDEM's architecture is portrayed in figure 1.

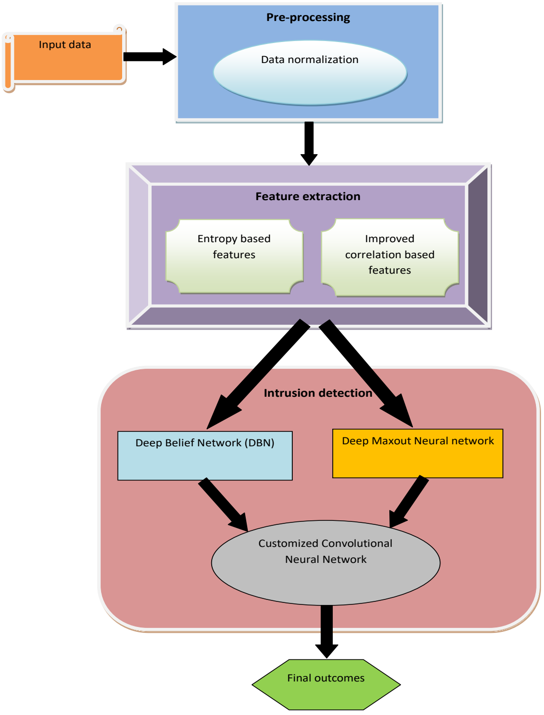

Input data
Pre-processing
Data normalization
Feature extraction
Entropy based features
Improved correlation based features
Intrusion detection
Deep Belief Network (DBN)
Deep Maxout Neural network
Customized Convolutional Neural Network
Final outcomes

Fig 1: Architecture of proposed IDSDEM

3.1. Pre-processing

Initially the input data I_d having the size of (18000×3) is sent to the pre-processing stage as input. The technique of transforming unclean data into clean data is known as data pre-processing. Before performing the algorithm, the dataset gets pre-processed to search for noisy data, missing values as well as other abnormalities. We have utilized data normalization process in this pre-processing stage. The process of arranging a relational data in accordance with a set of normal forms in attempt to eliminate data redundancy but also increase data integrity is known as data normalization [27]. After pre-processing this pre-processed data P_d get subjected to feature extraction stage which is explained below.

3.2. Feature extraction

Feature extraction is the second stage of our work. The technique of turning raw data into numeric features which could be handled while keeping the data in the initial data set is known as feature extraction. Compared to using AI techniques on the raw data directly, it produces superior outcomes. Here the features like entropy based and improved correlation based features were extracted from the pre-processed data P_d .

3.2.1. Entropy based feature

Entropy measures the disorder or impurity in the data that artificial intelligence algorithms process [28]. Considering a data set containing n classes; the entropy H may be calculated using the following formula (eqn. 1):

$$H\left(P_d\right)=-\sum_{a=1}^{n}\Pr\left(P_d\right)*\log_2\Pr\left(P_d\right)$$

(1)

Here P_d denotes the data while $\Pr$ denotes the probability of data.

3.2.2. Improved Correlation based feature

Correlation is indeed a process that establishes correlations among two variables. Correlation methods describe the relation among 2 variables in a single number known as the correlation coefficient. The correlation coefficient is varies from -1 to +1.

Correlation is assessed using different kinds of correlation coefficients depending on the characteristics of the compared data. The Pearson coefficient is among the most prevalent, and it measures the direction and strength of a linear correlation among two variables [29]. The numerical formula for Pearson coefficient is given in eqn. (2).

$$PC=\frac{\sum(b-mb)(c-mc)}{\sqrt{\sum(b-mb)^2\sum(c-mc)^2}}$$

(2)

The improved correlation uses the following eqn. (3), instead of the conventional correlation coefficient eqn. (2).

$$PC=\frac{\sum(b-mb)(c-mc)}{\sqrt{\sum(b-mb)^2\sum(c-mc)^2}}*CV$$

(3)

Here PC symbolizes the correlation coefficient. In a sample b and c values were denoted as b and c. Also b and c value's means were denoted as mb and mc. In eqn. (3), CV

function is evaluated using coefficient of variation (CV) formula (eqn. 4). Here α and δ were the data’s standard deviation and mean.

$$CV=\frac{\alpha}{\delta} \qquad (4)$$

These extracted features $F_{ex}=\left[H(y)\quad PC\right]$ were subjected to final intrusion detection stage where ensemble classification process takes place.

3.3. Intrusion detection

In this final intrusion detection stage, the stacked ensemble model with the classifiers like DBN; Deep maxout and CCNN are utilized to provide an accurate and efficient outcomes. Firstly, extracted feature F_{ex} having the size of 18000×3 is sent to DBN and Deep maxout networks as the input. In our work, the classification outcomes from DBN and Deep maxout were considered as features for further process. The classification outcomes (features) were given to CCNN as input to find an accurate outcome.

3.3.1. DBN

A DBN is made up of two kinds of neural networks (NNs): Belief Networks as well as Restricted Boltzmann Machines (RBMs). DBN is composed of a series of RBMs. Each RBM's hidden layer has been linked to the subsequent RBM's visual input layer. Every hidden layer of a NN can learn to express features in the initial input data which obtain higher-order correlations [30].

The RBM is indeed an undirected visual system containing visible units $VU\in\{0,1\}^f$ as well as hidden units $HU\in\{0,1\}^D$, with every visible unit linked to the corresponding hidden unit. The model may be described as an energy function for a specific set of (VU,HU) values as given in eqn. (5):

$$E(VU,HU)=-\sum_{i\in visible}d_iVU_i-\sum_{j\in hidden}e_jHU_j-\sum_{i,j}VU_iHU_jw_{ij}$$

(5)

Where hidden unit j and visible unit i’s binary states were symbolized as HU_j and VU_i while HU_j and VU_j’s biases were denoted as e_j and d_i Also w_{ij} denotes the weight among HU_j and VU_i. The joint distribution over VU and HU has been defined as given in eqn.(6):

$$A(VU,HU)=\frac{1}{Z}\exp\left(-E\left(VU,HU\right)\right),$$

(6)

The preceding formulas (eqn. (7) and eqn. (8)) make it simple to acquire an unbiased sample for the configuration of visible as well as hidden units. The energy function is written as in eqn. (9).

$$p\left(VU_i = 1 \mid HU\right) = \chi\left(d_i + \sum_j HU_j w_{ij}\right),$$

(7)

$$p\left(HU_j = 1 \mid VU\right) = \chi\left(e_j + \sum_i VU_i w_{ij}\right)$$

(8)

$$E\left(VU, HU\right) = -\sum_{i \in visible} \frac{\left(VU_i - d_i\right)^2}{2\chi_i^2} VU_i - \sum_{j \in hidden} e_j HU_j - \sum_{i,j} \frac{VU_i}{\chi_i} HU_j w_{ij}$$

(9)

Where χ_i symbolizes the standard deviation. The technique for obtaining an unbiased sample of the hidden unit's state remains unchanged; however the visible units were created utilizing the following conditional distribution as in eqn. (10):

$$p\left(VU_i = 1 \mid HU\right) = N\left(e_i + \chi_i \sum_{j=1}^{f} w_{ij} HU_j, \chi_i^2\right),$$

(10)

For all RBM forms, feature extraction depending on maximum-probability learning remains unsolvable. As a result, efficient learning may be accomplished by attempting to approximate the contrastive divergence objective's gradient.

The key concept for DBN training is as follows: An RBM learn the model parameter θ that could be specified as $p\left(VU_i = 1 \mid HU\right)$ and $p\left(HU_j = 1 \mid VU\right)$. After learning θ, the hidden activity vectors generated from the training data is obtained by $p\left(HU_j = 1 \mid VU\right)$. Afterwards another RBM is trained by the utilization of hidden activity vectors. Then, back propagation is employed to train the whole DBN in order to achieve the best biases and weights for intrusion detection. Sigmoid activation function is utilized in our work.

Finally, one of the classes from 0→black hole, 1→flooding, 2→grey hole, 3→normal and 4→TDMA were given as the outcome. This classification outcome is considered as a feature in our work, which is given to CCNN as input along with other features.

3.3.2. Deep Maxout

Every neuron in a maxout neural network seems to have a group of C candidate pieces. The neuron activation has been determined by taking the maximum value across all C components. [31][32][33][34]. The l-th hidden layer's Q-th node is denoted as HD_l^Q, as well as its equivalent pieces are denoted as X_l^{QJ}. The relationship among them satisfies the following eqn. (11)[35].

$$HD_l^Q = \max_{J \in 1,\dots C} X_l^{QJ}$$

(12)

Here X_l^{QJ} is derived from the layer below via forward propagation, i.e. given in eqn. (13).

$$X_l = W_{l-1}^T HD_{l-1} + g_l$$

(13)

Where $X_l \in R^X$ denotes the l-th layer's vector to be max pooled, the components of which include X_l^{QJ} The $l-1$th layer's maxout activation vector as well as its weight matrix were symbolized as $HD_{l-1} \in R^{HD}$ and $W_{l-1} \in R^{HD \times X}$, while the l-th layer's bias vector is denoted by $g_l \in R^X$ [36].

Maxout network's forward-propagation process is identical as other feed-forward NNs, except the activation computation (eqn. (3) and (4)). The forward-propagation process of the maxout network is the same as other feed-forward neural networks except that the activation computation follows equation (3) and (4). During the training phase, the gradient for every maxout neuron is always 1, while only the weights associated to the piece with the highest activation in every group $\left\{X_l^{QJ}\right\} J \in 1 \dots C$ for $l \in \left[1, V\right]$ and $i \in \left[1, N^l\right]$ were updated. In our deep maxout model, we use an optimization approach to repeatedly optimize the objective function. We use the Adam optimization approach to optimize our deep maxout model. There are numerous advantages of utilising Adam optimizer. T his is computationally efficient and easy to execute. Moreover, it is an empirically proven optimization strategy for deep neural networks. The Adam optimizer employs element-wise squared gradients, as well as adaptive 1st & 2nd moments[37].

$$BI_{M1}\left(t\right) = \beta_1 . BI_{M1}\left(t-1\right) + \left(1-\beta_1\right) . G_L\left(t\right)$$

(14)

$$BI_{M2}\left(t\right) = \beta_2 . BI_{M2}\left(t-1\right) + \left(1-\beta_2\right) . G_L^2\left(t\right)$$

(15)

$$u_{M1}\left(t\right) = \frac{BI_{M1}\left(t\right)}{1-\beta_1^t}$$

(16)

$$u_{M2}(t) = \frac{BI_{M2}(t)}{1-\beta_2^t}$$

(17)

Consider BI_{M1}, BI_{M2}, u_{M1}, u_{M2}, η, β_1, β_2 and G_L represent the biased 1st moment, biased 2nd moment, unbiased 1st moment, unbiased 2nd moment, learning rate, first moment's exponential decay rate, second moment's exponential decay rate, gradient correspondingly. The Adam optimization approach may be mathematically expressed utilizing Eqs. 14, 15, 16, and 17. Sequentially, a categorical cross-entropy loss function has been used to minimize the training loss. Finally, one of the classes from 0→black hole, 1→flooding, 2→grey hole, 3→normal and 4→TDMA were given as the outcome. This classification outcome is considered as a feature in our work, which is given to CCNN as input along with other features and DBN outcomes[36].

3.3.3. Customized CNN

The feature size of 18000×5 is given to CCNN as input. This includes the input features and the features from the outcome of DBN and Deep maxout[37]. Our CCNN is conceptually and architecturally simple. This CCNN is designed as a sequential design with six convolutional blocks as well as two fully-connected layers. CNN's heart seems to be the convolution layer. Proposed CCNN's architecture has been given in figure 2.

To generate output feature maps, this layer performs convolution operations on the input feature map's tiny, localised regions[40][38]. In contrast to traditional NN, every neuron in a hidden layer gets coupled to a small portion of the preceding layer known as the receptive field. The convolution layer is made up of small-dimension learnable filters or kernels[39]. To generate output feature maps, such kernels concatenate with receptive fields and expand throughout the whole width, height, as well as depth of the input volume. Because kernel sizes determine weights, every hidden neuron contains a bias as well as weights equal to the kernel dimension associated to its receptive field [41]. Softmax activation is employed for every convolutional layer in all of the six convolutional blocks, as well as a max pooling layer gets coupled to the final convolutional layer in every block. A pooling layer is indeed an essential part of a NN that is employed to reduce the spatial dimensionality of the convolution layer' output, hence lowering the count of parameters as well as network computational complexity but also controlling over fitting. In our work, we used max pooling, which extracts a solitary maximal value from a neuron's cluster in the precursive layer.

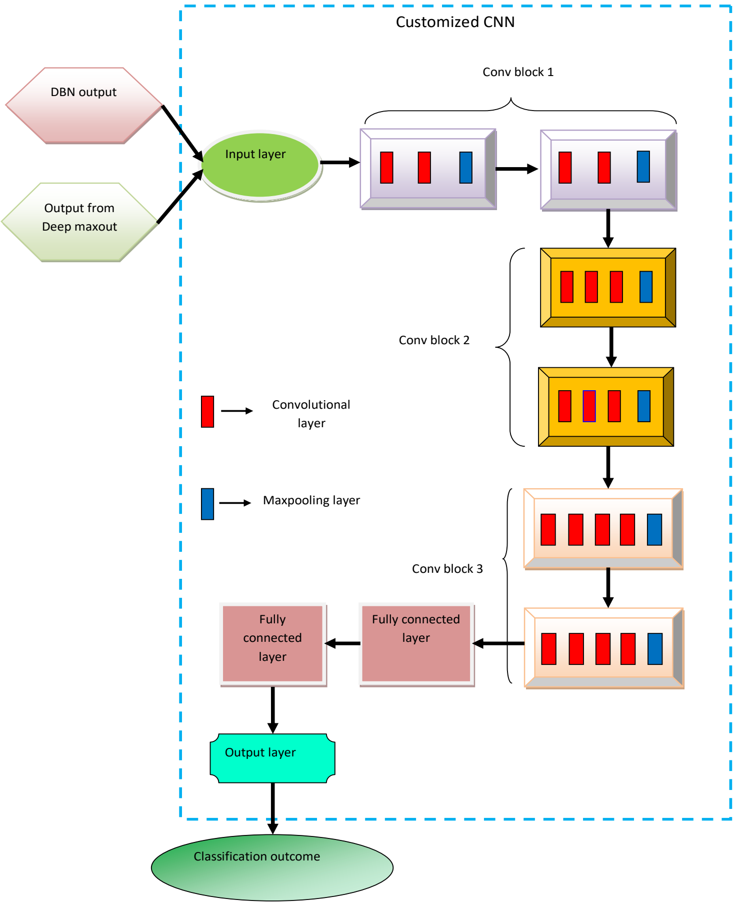


Fig 2: Architecture of CCNN

A fully connected layer's function is to merge information from pooling and convolution layers to get a probable class score for categorization of the data streams. The final fully-connected layer is activated using Softmax. This activation function outperforms others in classification because it compresses the outputs of every segment among 0 and 1. The loss function has been categorical entropy loss, and also the optimization is carried out using the Stochastic Gradient Decent algorithm (SGD). Final classification outcomes or output classes (0→black hole, 1→flooding, 2→grey hole, 3→normal and 4→TDMA) were given as the outcome.

4. Simulation Procedure

The proposed Intrusion Detection framework was implemented in PYTHON and the dataset was assembled in [34]. The Intrusion Detection by Stacked Deep Ensemble Model (IDSDEM) was determined over the following algorithms: Decision Tree (DT), Deep Neural Network (DNN), Random Forest (RF), Long Short Term Memory (LSTM), Recurrent Neural Network (RNN), Support Vector Machine (SVM) and Bidirectional Gated Recurrent unit (Bi-GRU), respectively. Further, the evaluation was carried out with respect to precision, FPR, accuracy, NPV and so on as well as the relevant outcomes are portrayed below.

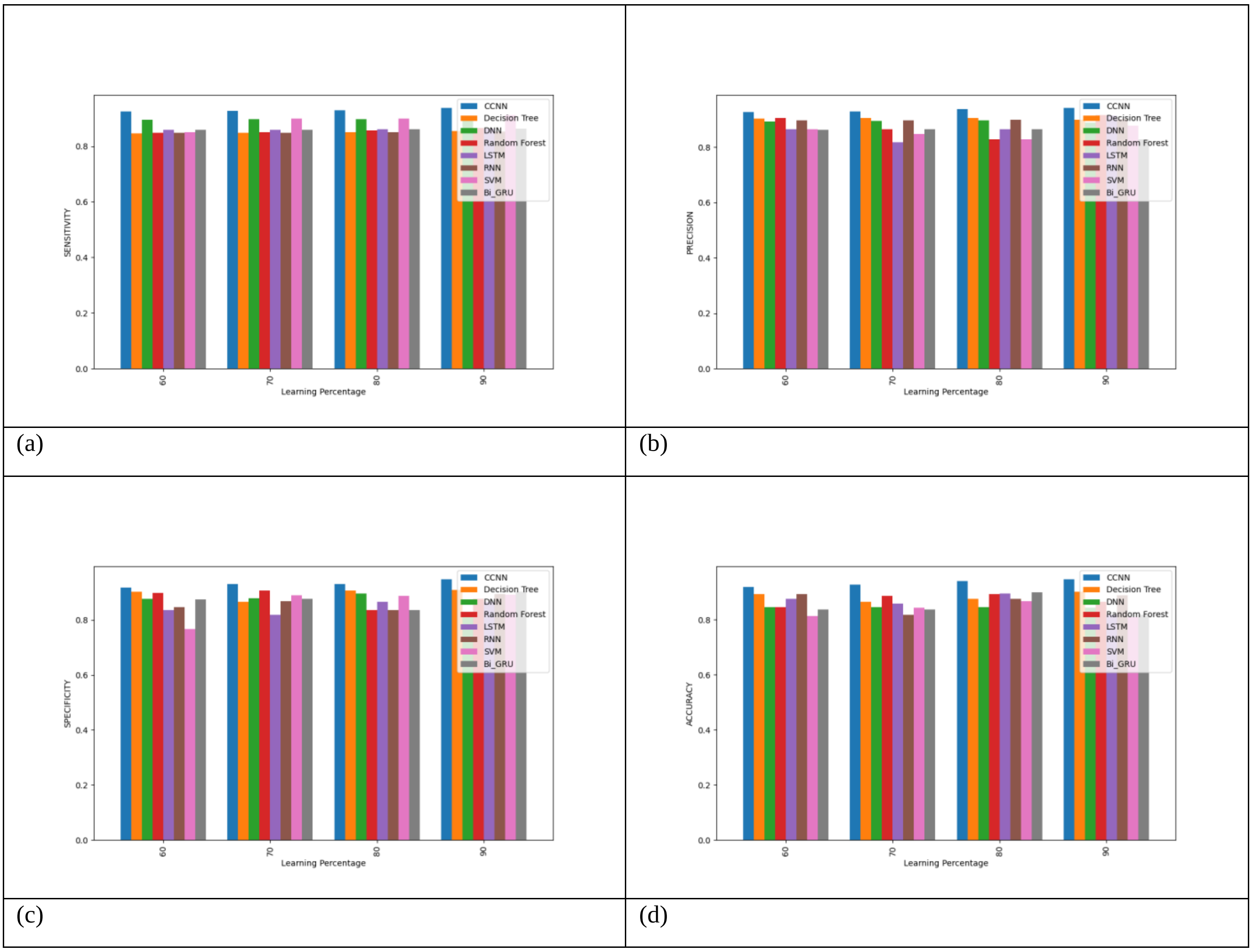
4.1 Dataset Description

“A specialized dataset for WSN is developed to help better detect and classify four types of Denial of Service (DoS) attacks: Blackhole, Grayhole, Flooding, and Scheduling attacks. This paper considers the use of LEACH protocol which is one of the most popular hierarchical routing protocols in WSNs. A scheme has been defined to collect data from Network Simulator 2 (NS-2) and then processed to produce 23 features. The collected dataset is called WSN-DS. Artificial Neural Network (ANN) has been trained on the dataset to detect and classify different DoS attacks. The results show that WSNDS improved the ability of IDS to achieve higher classification accuracy rate. WEKA toolbox was used with holdout and 10-Fold Cross Validation methods. The best results were achieved with 10-Fold Cross Validation with one hidden layer. The classification accuracies of attacks were 92.8%, 99.4%, 92.2%, 75.6%, and 99.8% for Blackhole, Flooding, Scheduling, and Grayhole attacks, in addition to the normal case (without attacks), respectively.”

4.2 Assessment on positive measure of the IDSDEM over the traditional methodologies by modifying the learning percentage

The interpretation of the IDSDEM is related to the current algorithms such as DT, DNN, RF, LSTM, RNN, SVM and Bi-GRU is depicted in fig 3. Additionally, it is verified by changing the learning percentage. For better intrusion detection performance, the positive measures should be higher. Likewise, the IDSDEM generated greater positive metric values over the other schemes. The IDSDEM for the 90th learning percentage recorded the highest sensitivity, reaching 94.87%, whilst the current schemes recorded the lowest sensitivity, namely, DT=82.64%, DNN=88.28%, RF=83.92%, LSTM=84.46%, RNN=81.34%, SVM=86.69% and Bi-GRU=85.73%, respectively. At the 60% of learning percentage, the approaches like DT, DNN, RF, LSTM, RNN, SVM and Bi-GRU attained the precision of 88.24%, 86.38%, 89.67%, 83.59%, 89.98%, 84.76% and 85.44%, although the IDSDEM acquired the precision of 91.46%.

Moreover, fig 3(c) and fig 3(d) demonstrate the results of comparing the specificity and accuracy measure of the IDSDEM and the extant strategies. Regarding the fig 3(c), the IDSDEM accomplished the maximum specificity of 96.98%, at the learning percentage 80, mean while the DT is 90.86%, DNN is 89.74%, RF is 78.19%, LSTM is 83.29%, RNN is 80.35%, SVM is 87.56% and Bi-GRU is 77.94%, respectively. While estimating the fig 3(d), the IDSDEM accomplished higher detection accuracy than the other current methods. While adjusting the learning percentage to 70%, the IDSDEM scored the accuracy of 94.77%, whilst the DT=86.34%, DNN=84.62%, RF=88.46%, LSTM=87.59%, RNN=78.66%, SVM=82.21% and Bi-GRU=81.49%, respectively. Furthermore, the IDSDEM scored the accuracy of 90.38% in the 60th learning percentage, though it recorded the maximal accuracy of 96.24% at the learning percentage 90%. Thus, the outcomes of this assessment expressed that the IDSDEM model’s positive measure rating is advantageous and this enhancement is due to the improved correlation based features and customized CNN.



(a)

(b)

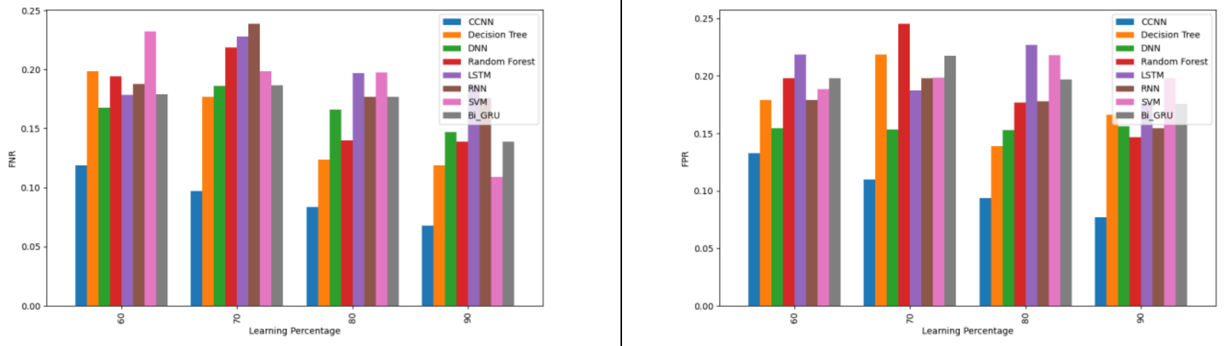
(c)

(d)

Fig 3: Evaluation on positive measure of the IDSDEM versus standard methods with regard to a) Sensitivity b) Precision c) Specificity d) Accuracy with adjusting the learning percentage

4.3 Assessment on FNR and FPR of the IDSDEM over the traditional methodologies by modifying the learning percentage

Fig 4 describes the FNR and FPR of the IDSDEM and the extant methods like DT, DNN, RF, LSTM, RNN, SVM and Bi-GRU. Low negative measures should be required for the enhanced intrusion detection system. More specifically, for the learning percentage 90, the FNR of the IDSDEM is 0.063, despite the fact that established schemes attained higher FNR, like, DT is 0.129, DNN is 0.148, RF is 0.136, LSTM is 0.187, RNN is 0.176, SVM is 0.122 and Bi-GRU is 0.138, respectively. The IDSDEM FPR ranges from 0.13 to 0.07, making it lowest among the DT, DNN, RF, LSTM, RNN, SVM, and Bi-GRU. Especially, when fixing the learning percentage to 80, the IDSDEM attained the FPR of 0.087, although the DT=0.143, DNN=0.148, RF=0.179, LSTM=0.236, RNN=0.181, SVM=0.215 and Bi-GRU=0.194, respectively. This estimation implies that the IDSDEM delineated the system is more adaptive for intrusion detection, which is because of the contributions we made in this work.

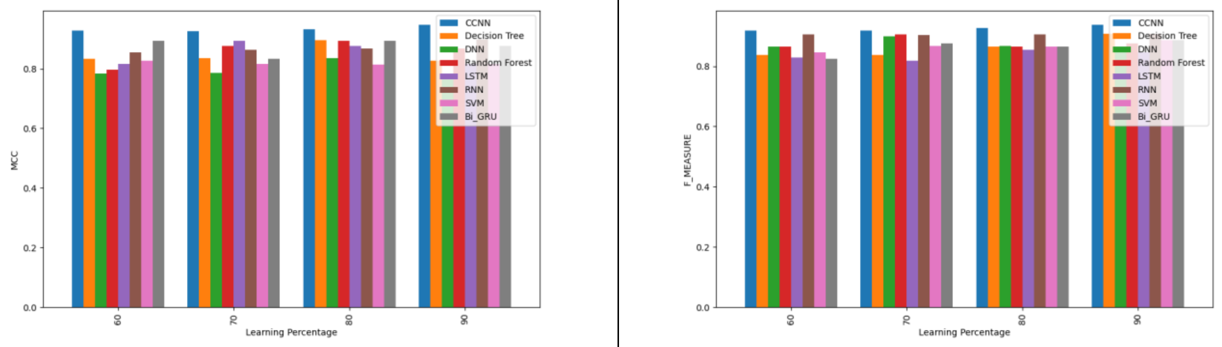


(a) (b)

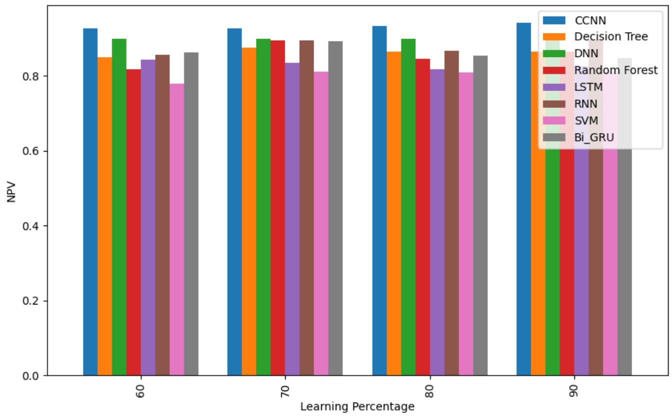
Fig 4 Evaluation on negative measure of the IDSDEM versus standard methods with regard to a) FNR b) FPR with adjusting the learning percentage

4.4 Assessment on other measure of the IDSDEM over the traditional methodologies by modifying the learning percentage

The other measure evaluation on IDSDEM is assessed over the DT, DNN, RF, LSTM, RNN, SVM and Bi-GRU by modifying the learning percentage (60 to 90). The compatible results are overviewed in fig 5. Actually, the other measures should be maximal for better intrusion detection performance. In that way, the IDSDEM is found to accomplish highest (F-measure, MCC and NPV) outcomes over compared models like DT, DNN, RF, LSTM, RNN, SVM and Bi-GRU. The MCC of the IDSDEM is 97.69% (Learning Percentage=90%), this is superior to DT=78.85%, DNN=75.92%, RF=82.46%, LSTM=79.32%, RNN=84.76%, SVM=76.59% and Bi-GRU=79.74%, respectively. While exploring the fig 5(b), the IDSDEM acquired the greatest F-measure of 93.57% in the 70% of learning percentage, though the DT, DNN, RF, LSTM, RNN, SVM and Bi-GRU gained the lowest F-measure of 80.98%, 81.54%, 85.28%, 86.97%, 89.78%, 87.23% and 85.42%, respectively. Furthermore, for the 80th learning percentage, the IDSDEM obtained the NPV of 81.78%, which is preferable than DT (81.78%), DNN (84.91%), RF (79.29%), LSTM (77.67%), RNN (83.54%), SVM (78.65%) and Bi-GRU (80.24%), respectively. As a consequence, this assessment affirms that the IDSDEM accomplish accurate intrusion detection as if contrasted to the extant algorithms.



(a) (b)



(c)

Fig 5: Evaluation on negative measure of the IDSDEM versus standard methods with regard to a) MCC b) F-measure c) NPV with adjusting the learning percentage

4.5 Statistical examination on IDSDEM over the current techniques in terms of accuracy for varied case scenarios

The statistical evaluation was carried out by utilizing the Mean, Standard Deviation, Minimum, Median and Maximum case scenario. Subsequently, the IDSDEM is compared with DT, DNN, RF, LSTM, RNN, SVM and Bi-GRU as well as the findings are presented in table II. For better intrusion detection system, the detection accuracy must be higher. Similarly, the IDSDEM attained the higher detection accuracy than the established schemes in all the case scenarios. As per the minimum case scenario, the IDSDEM generated the accuracy of 0.91893, in contrast the DT=0.865462, DNN=0.844095, RF=0.844095, LSTM=0.818686, RNN=0.818568, SVM=0.809286 and Bi-GRU=0.818684, respectively. Similarly, the IDSDEM accomplished the accuracy of 0.934276 under the median case scenario, although, the current approaches recorded the lowest accuracy rate, including, DT is 0.885116, DNN is 0.845994, RF is 0.876569, LSTM is 0.867202, RNN is 0.883156, SVM is 0.828219 and Bi-GRU is 0.837807, respectively. Therefore, the improved performance of the IDSDEM has maximized the accuracy, this accentuate that IDSDEM can bestow trustworthy results for Intrusion Detection.

Table II: Statistical Assessment on IDSDEM over the standard algorithms in terms of accuracy for Mean, Standard Deviation, Minimum, Median and Maximum case scenario

	IDSDEM	DT	DNN	RF	LSTM	RNN	SVM	Bi-GRU
Mean	0.933705	0.884397	0.845756	0.873346	0.862416	0.869661	0.833315	0.848646
Standard Deviation	0.01078	0.01424	0.001076	0.018691	0.028578	0.030169	0.023859	0.03082
Minimum	0.918983	0.865462	0.844095	0.846573	0.818686	0.818568	0.809286	0.818684
Median	0.934276	0.885116	0.845994	0.876569	0.867207	0.883156	0.828219	0.837807
Maximum	0.947286	0.901873	0.846938	0.893674	0.896564	0.893764	0.867537	0.900287

4.6 Ablation analysis on IDSDEM, model without improved correlation and model without feature extraction for positive, negative and other measures

The ablation evaluation on IDSDEM, model without improved correlation and model without feature extraction for various types of measures are described in table III. The MCC of the IDSDEM is 92.67%, model without improved correlation is

81.87% and model without feature extraction is 92.67%. The IDSDEM, model without improved correlation and model without feature extraction has obtained the accuracy of 85.76%, 89.36% and 92.76%, respectively. The F-measure, FNR and Precision of the IDSDEM is 91.86%, 0.097363 and 92.86%, respectively. This superiority is because of the customized CNN and Improved correlation based features we have utilized in this work.

Table III: Ablation evaluation on IDSDEM, model without improved correlation and model without feature extraction with regard to positive, negative and other measures

	Model without Improved Correlation	Model without feature extraction	IDSDEM
Accuracy	0.857683	0.893663	0.927658
MCC	0.818787	0.859875	0.926757
Precision	0.798386	0.841472	0.928683
F-measure	0.827879	0.873478	0.918682
Specificity	0.900078	0.889244	0.928663
NPV	0.898738	0.802874	0.927569
FPR	0.186723	0.110756	0.109873
Sensitivity	0.867363	0.838685	0.929797
FNR	0.17773	0.198783	0.097363

4.7 Classifier Comparison on IDSDEM is compared to the conventional methodologies for different measures by modifying the learning percentage

The classifier comparison on IDSDEM is contradicted over the DT, DNN, RF, LSTM, RNN, SVM and Bi-GRU for positive, negative and other measures by varying the learning percentage from 60 to 90 is represented in table IV. Here, the IDSDEM attained highest positive and other measures as well as it acquired lowest negative measures. For the 70th learning percentage, the FNR of the IDSDEM is 0.097363, though the extant methods scored maximal FNR, notably, DT=0.17672, DNN=0.18627, RF=0.218739, LSTM=0.227635, RNN=0.238684, SVM=0.198737 and Bi-GRU=0.186379, respectively. The NPV of the IDSDEM for the 60th learning percentage is 92.75%, mean while the DT is 84.98%, DNN is 89.87%, RF is 81.75%, LSTM is 84.36%, RNN is 85.67%, SVM is 77.92% and Bi-GRU is 86.34%, respectively. While adjusting the learning percentage to 70%, the sensitivity, FPR and MCC of the IDSDEM is 92.76%, 0.109873 and 92.67%. This signified the potentiality of the IDSDEM for intrusion detection.

Table IV: Classifier comparison on IDSDEM over the established classifiers in terms of varied measures by adjusting the learning percentage

Learning Percentage=60%								
	IDSDEM	DT	DNN	RF	LSTM	RNN	SVM	Bi-GRU
Sensitivity	0.925296	0.846488	0.895293	0.849169	0.858538	0.848064	0.850737	0.858538
MCC	0.927687	0.832764	0.784237	0.796373	0.815463	0.854532	0.827576	0.893674
Precision	0.927764	0.903516	0.89306	0.905529	0.865657	0.897822	0.864532	0.863667
F-measure	0.917673	0.837654	0.865643	0.865764	0.828635	0.904693	0.846543	0.825673
Accuracy	0.918983	0.893664	0.845568	0.846573	0.87565	0.893764	0.812784	0.837874
NPV	0.927584	0.84987	0.898721	0.817573	0.843657	0.856783	0.779229	0.863465
FPR	0.132863	0.178783	0.154432	0.197873	0.218687	0.17904	0.188368	0.19793

Specificity	0.917683	0.903516	0.87645	0.898657	0.836576	0.846753	0.767976	0.875634
FNR	0.118735	0.198787	0.167699	0.193987	0.178674	0.187894	0.232024	0.178793
Learning Percentage=70%								
	IDSDEM	DT	DNN	RF	LSTM	RNN	SVM	Bi-GRU
Sensitivity	0.927658	0.848363	0.897029	0.850542	0.858059	0.847555	0.900028	0.858059
MCC	0.926757	0.83657	0.78705	0.87682	0.893876	0.86463	0.815213	0.834565
Precision	0.928683	0.904792	0.895566	0.864632	0.816753	0.897454	0.847693	0.865766
F-measure	0.918682	0.836547	0.898624	0.906457	0.817583	0.904347	0.86773	0.875763
Accuracy	0.928663	0.865462	0.84642	0.887673	0.858763	0.818568	0.843655	0.83774
NPV	0.927569	0.876537	0.899314	0.895756	0.834246	0.89446	0.810861	0.893676
FPR	0.109873	0.21863	0.15358	0.245364	0.18725	0.197827	0.198674	0.217638
Specificity	0.929797	0.865632	0.878686	0.906457	0.817657	0.867546	0.889436	0.87673
FNR	0.097363	0.17672	0.18627	0.218739	0.227635	0.238684	0.198737	0.186379
Learning Percentage=80%								
	IDSDEM	DT	DNN	RF	LSTM	RNN	SVM	Bi-GRU
Sensitivity	0.929288	0.850591	0.898067	0.856779	0.860363	0.85	0.900288	0.860363
MCC	0.932877	0.896757	0.83634	0.893776	0.876764	0.867564	0.813768	0.893567
Precision	0.937682	0.906304	0.897049	0.828734	0.865673	0.899218	0.827657	0.86462
F-measure	0.927535	0.865766	0.86753	0.86453	0.853467	0.906005	0.864678	0.865673
Accuracy	0.939888	0.876567	0.846938	0.893674	0.896564	0.87664	0.867537	0.900287
NPV	0.932975	0.86573	0.899675	0.846462	0.816757	0.866543	0.810125	0.854653
FPR	0.093764	0.138979	0.153062	0.176564	0.226757	0.177653	0.217873	0.196763
Specificity	0.93108	0.906304	0.896733	0.836574	0.865467	0.836573	0.887847	0.836758
FNR	0.083657	0.123768	0.165724	0.139794	0.196769	0.176573	0.197688	0.176764
Learning Percentage=90%								
	IDSDEM	DT	DNN	RF	LSTM	RNN	SVM	Bi-GRU
Sensitivity	0.937657	0.853842	0.892197	0.865975	0.86267	0.852449	0.91238	0.86267
MCC	0.947658	0.827688	0.77918	0.867577	0.816753	0.898737	0.816771	0.875653
Precision	0.941869	0.89893	0.888501	0.916796	0.914441	0.90098	0.877566	0.827678
F-measure	0.937568	0.908505	0.864652	0.87672	0.846577	0.907661	0.886754	0.887365
Accuracy	0.947286	0.901873	0.844095	0.865465	0.818686	0.889673	0.809286	0.818684
NPV	0.941979	0.865463	0.897693	0.865464	0.826573	0.896757	0.811667	0.846757
FPR	0.076757	0.1663	0.155905	0.146538	0.176384	0.15458	0.197874	0.175653
Specificity	0.94798	0.908505	0.818673	0.876567	0.86563	0.893277	0.891137	0.901987
FNR	0.067538	0.118768	0.146758	0.138624	0.187893	0.175789	0.108863	0.138679

5. Conclusion

Because of the rise of the unlimited communication network as well as the rising amount of networked digital gadgets in current years,, there has been a growing worry about cyber security, which attempts to maintain either the system's data or communication technology. Intruders develop new attack kinds on a regular basis; consequently, in order to prevent these assaults, they must first be accurately identified by the IDSs in use, and then appropriate responses must be provided. Novel IDS named IDSDEM, which has three working stages. Initially pre-processing process takes place, which utilizes the data normalization to provide a consistent data for further processing. Afterwards feature extraction stage takes place where the features like entropy based as well as improved correlation based features get extracted. Finally, intrusion detection was conducted where a stacked deep ensemble model which includes the classifiers like Deep Belief Network (DBN), Deep Maxout Network and Customized Convolutional Neural Network (CCNN) was utilized to provide effective intrusion detection. From the outcomes it's proved that our IDSDEM can offer superior performance in the intrusion detection task than other extant approaches. The Accuracy, Sensitivity, specificity and Precision of the IDSDEM is 94.77%, 94.87%, 96.98% and 91.46%, respectively.

Reference

[1] Liu, G., Zhao, H., Fan, F., Liu, G., Xu, Q. and Nazir, S., "An Enhanced Intrusion Detection Model Based on Improved kNN in WSNs". Sensors, 22(4), p.1407, (2022).

[2] Maheswari, M. and Karthika, R.A., "A novel QoS based secure unequal clustering protocol with intrusion detection system in wireless sensor networks". Wireless Personal Communications, 118(2), pp.1535-1557, (2021).

[3] Gowdhaman, V. and Dhanapal, R., "An intrusion detection system for wireless sensor networks using deep neural network". Soft Computing, 26(23), pp.13059-13067, (2022).

[4] Elsaid, S.A. and Albatati, N.S., "An optimized collaborative intrusion detection system for wireless sensor networks". Soft Computing, 24(16), pp.12553-12567, (2020).

[5] Mohapatra, H., Rath, S., Panda, S. and Kumar, R., "Handling of man-in-the-middle attack in wsn through intrusion detection system". International journal, 8(5), pp.1503-1510, (2020).

[6] Safaldin, M., Otair, M. and Abualigah, L., "Improved binary gray wolf optimizer and SVM for intrusion detection system in wireless sensor networks". Journal of ambient intelligence and humanized computing, 12(2), pp.1559-1576, (2021).

[7] Zhang, W., Han, D., Li, K.C. and Massetto, F.I., "Wireless sensor network intrusion detection system based on MK-ELM". Soft Computing, 24(16), pp.12361-12374, (2020).

[8] Gavel, S., Raghuvanshi, A.S. and Tiwari, S., "A novel density estimation based intrusion detection technique with Pearson's divergence for wireless sensor networks". ISA transactions, 111, pp.180-191, (2021).

[9] Godala, S. and Vaddella, R.P.V., "A study on intrusion detection system in wireless sensor networks". International Journal of Communication Networks and Information Security, 12(1), pp.127-141, (2020).

[10] Baraneetharan, E., "Role of machine learning algorithms intrusion detection in WSNs: a survey". Journal of Information Technology, 2(03), pp.161-173, (2020).

[11] Gite, P., Chouhan, K., Krishna, K.M., Nayak, C.K., Soni, M. and Shrivastava, A., "ML Based Intrusion Detection Scheme for various types of attacks in a WSN using C4. 5 and CART classifiers". Materials Today: Proceedings, (2021).

[12] Singh, A., Nagar, J., Sharma, S. and Kotiyal, V., "A Gaussian process regression approach to predict the k-barrier coverage probability for intrusion detection in wireless sensor networks". Expert Systems with Applications, 172, p.114603, (2021).

[13] Aldweesh, A., Derhab, A. and Emam, A.Z., 2020. Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues. Knowledge-Based Systems, 189, p.105124.

[14] Sinha, S. and Paul, A., "Neuro-fuzzy based intrusion detection system for wireless sensor network". Wireless Personal Communications, 114(1), pp.835-851, (2020).

[15] Nancy, P., Muthurajkumar, S., Ganapathy, S., Santhosh Kumar, S.V.N., Selvi, M. and Arputharaj, K., "Intrusion detection using dynamic feature selection and fuzzy temporal decision tree classification for wireless sensor networks". IET Communications, 14(5), pp.888-895, (2020).

[16] Almomani, I. and Alromi, A., "Integrating software engineering processes in the development of efficient intrusion detection systems in wireless sensor networks. Sensors, 20(5)", p.1375, (2020).

[17] Amaran, S. and Mohan, R.M., "Intrusion detection system using optimal support vector machine for wireless sensor networks". In 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) (pp. 1100-1104). IEEE, (2021).

[18] Kaur, N. and Rattan, P., "A critical review of intrusion detection systems in WSN: challenges & future directions". Annals of the Romanian Society for Cell Biology, pp.3020-3028, (2021).

[19] Alruhaily, N.M. and Ibrahim, D.M., “A multi-layer machine learning-based intrusion detection system for wireless sensor networks”. International Journal of Advanced Computer Science and Applications, 12(4), pp.281-288, (2021).

[20] Umarani, C. and Kannan, S., “Intrusion detection system using hybrid tissue growing algorithm for wireless sensor network”. Peer-to-Peer Networking and Applications, 13(3), pp.752-761, (2020).

[21] Otoum, S., Kantarci, B. and Mouftah, H.T., “A novel ensemble method for advanced intrusion detection in wireless sensor networks”. In Icc 2020-2020 ieee international conference on communications (icc) (pp. 1-6). IEEE, (2020).

[22] Jiang, S., Zhao, J. and Xu, X., “SLGBM: An intrusion detection mechanism for wireless sensor networks in smart environments”. IEEE Access, 8, pp.169548-169558, (2020).

[23] Alsahli, M.S., Almasri, M.M., Al-Akhras, M., Al-Issa, A.I. and Alawairdhi, M., “Evaluation of machine learning algorithms for intrusion detection system in WSN”. International Journal of Advanced Computer Science and Applications, 12(5), (2021).

[24] Alwan, M.H., Hammadi, Y.I., Mahmood, O.A., Muthanna, A. and Koucheryavy, A., “High Density Sensor Networks Intrusion Detection System for Anomaly Intruders Using the Slime Mould Algorithm”. Electronics, 11(20), p.3332, (2022).

[25] Abhale, A.B. and Manivannan, S.S., “Supervised machine learning classification algorithmic approach for finding anomaly type of intrusion detection in wireless sensor network”. Optical Memory and Neural Networks, 29(3), pp.244-256, (2020).

[26] Aljanabi, M., Ismail, M.A. and Ali, A.H., “Intrusion detection systems, issues, challenges, and needs”. International Journal of Computational Intelligence Systems, 14(1), pp.560-571, (2021).

[27] https://en.wikipedia.org/wiki/Database_normalization

[28] https://www.javatpoint.com/entropy-in-machine-learning

[29] https://www.google.com/search?q=pearson+correlation+formula&rlz=1C1CHBF

[30] Zhao, L., Wang, Z., Wang, X. and Liu, Q., “Driver drowsiness detection using facial dynamic fusion information and a DBN”. IET Intelligent Transport Systems, 12(2), pp.127-133, (2018).

[31] Cai, M., Shi, Y. and Liu, J., “Deep maxout neural networks for speech recognition”. In 2013 IEEE Workshop on Automatic Speech Recognition and Understanding (pp. 291-296). IEEE, (2013).

[32] Davanam, G., Kumar, T. P., & Kumar, M. S. (2021). Efficient energy management for reducing cross layer attacks in cognitive radio networks. Journal of Green Engineering, 11, 1412-1426.

[33] Kumar, M. S., Siddardha, B., Reddy, A. H., Reddy, C. V. S., Shaik, A. B., & Ganesh, D. (2022). APPLYING THE MODULAR ENCRYPTION STANDARD TO MOBILE CLOUD COMPUTING TO IMPROVE THE SAFETY OF HEALTH DATA. Journal of Pharmaceutical Negative Results, 1911-1917.

[34] P. Sai Kiran, "Power aware virtual machine placement in IaaS cloud using discrete firefly algorithm." Applied Nanoscience (2022): 1-9.

[35] Pavan Kumar, T.,(2021). Novel Defense Framework for Cross-layer Attacks in Cognitive Radio Networks. In International Conference on Intelligent and Smart Computing in Data Analytics (pp. 23-33). Springer, Singapore.

[36] Neelima, P., & Kumar, M. S. (2017, April). A memetic algorithm for multi objective vehicle routing problem with time windows. In 2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE) (pp. 1-8). IEEE.

[37] P. Sai Kiran, and M. S Kumar. "Resource aware virtual machine placement in IaaS cloud using bio-inspired firefly algorithm." Journal of Green Engineering 10 (2020): 9315-9327.

[38] Thummala Pavan Kumar, "Optimised Levenshtein centroid cross-layer defence for multi-hop cognitive radio networks." IET Communications 15.2 (2021): 245-256.

[39] Bari, A.H. and Gavrilova, M.L., “Novel Multi-layer Perceptron Architecture for Gait Recognition” (2019).

[40] Guo, P., Xue, Z., Mtema, Z., Yeates, K., Ginsburg, O., Demarco, M., Long, L.R., Schiffman, M. and Antani, S., “Ensemble deep learning for cervix image selection toward improving reliability in automated cervical precancer screening”. Diagnostics, 10(7), p.451, (2020).

[41] Almomani, I., Al-Kasasbeh, B. and Al-Akhras, M., 2016. WSN-DS: A dataset for intrusion detection systems in wireless sensor networks. Journal of Sensors, 2016.