International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2021/210725 Volume 8, Issue 6, November – December (2021)

RESEARCH ARTICLE

# Hybrid Intrusion Detection Method Based on Constraints Optimized SAE and Grid Search Based SVM-RBF on Cloud

Nirmalajyothi Narisetty
Department of Computer Science and Engineering, Acharya Nagarjuna University, Guntur, Andhra Pradesh, India
nirmala.narisetty@gmail.com

Gangadhara Rao Kancherla
Department of Computer Science and Engineering, Acharya Nagarjuna University, Guntur, Andhra Pradesh, India
kancherla123@gmail.com

Basaveswararao Bobba
Department of Computer Science and Engineering, Acharya Nagarjuna University, Guntur, Andhra Pradesh, India
bbrao@anu.ac.in

K. Swathi
Department of Computer Science and Engineering, K L Deemed to be University, Vaddeswaram, Guntur, Andhra Pradesh, India
dr.kswathi@kluniversity.in

Received: 01 October 2021 / Revised: 14 November 2021 / Accepted: 20 November 2021 / Published: 30 December 2021

**Abstract – The present era is facing lot of Security, Privacy, and Integrity issues because of tremendous development in communication technology, data storage devices, and computing advancements leading to unavoidable losses. As a result of the aforementioned technological revolutions day by day, many of the organizations or institutions started migrating to cloud environment. Because of this, security issues have increased coupled with the advent of new ways of penetration into networks. Unauthorized users and many professionals with malicious intent started exploiting the legitimate users through cyber-crimes. So, there is a need to implement a proper Intrusion Detection System with optimization procedures. This paper proposes a hybrid Intrusion Detection approach with a combination of Constraints Optimized Stacked Autoencoder (COSAE) for dimension reduction and grid search based SVM-RBF classifier (GSVM-RBF). The COSAE+GSVM-RBF model enhanced the performance using a two-fold. i) The SAE is optimized through regularization techniques with the adoption of weight and dropout constraints, ii) To enhance the performance of the SVM classifier with RBF for tuning the hyperparameters using grid search. Various experiments are conducted to validate this model with four activation functions Scaled Exponential Linear Unit (SELU), Rectified Linear Unit, softplus, and Exponential Linear Unit (ELU) for dimension reduction using COSAE. The improvements carried out in this paper result in exploding gradients and vanishing gradients avoids overfitting in large datasets, intrusion detection rate, gain in computational time, and 100% F-Measure in classifying minor class labels. The proposed approach is validated on the CICIDS2017 dataset. Further, a comparative analysis of the proposed approach with state-of-the-art approaches has been conducted. Based on the experimental results it is observed that the proposed approach outperforms the prevailing approaches.**

**Index Terms – Cloud Computing, Intrusion Detection, Stacked Autoencoder, Support Vector Machine, Regularization Constraints.**

## 1. INTRODUCTION

Cloud computing offers on-demand services, resources, and applications through the internet to industries, academia, and individuals with service delivery based on pay-as-you-use. These services are dynamically provisioned to users based on Service Level Agreement (SLA) established between Cloud Service Provider (CSP) and users. Ever since the advent of cloud computing the usage has gone up exponentially which would obviously increase the traffic and poses security threats as content, service, and many others like these are made available in the cloud. The frequency and potency of attacks in cloud environments necessitate the constant development of improved detection techniques. Among these, Distributed Denial of Service (DDoS) attacks is a type of Denial of Service (DoS) attack which disrupts the normal functioning of the cloud server ultimately service to legitimate users is denied. The interference of DDoS may directly or indirectly lead to financial loss or reputation loss of cloud service providers as well as users. Particularly this scenario in the cloud leads to EDoS attacks [1]. So, these kinds of attacks need to be addressed by providing a potential Network Intrusion Detection Systems (NIDS) for defending against security threats in the cloud environment. The main objective of NIDS is to monitor and classify malicious behavior of the network traffic to mitigate the consequences of attacks.

The statistical analysis-based NIDS models have become obsolete because of the increase in the volume of network flow data, there is an imminent need to bring intelligence into different kinds of NIDS techniques. Machine learning techniques have become so prominent to achieve this. Feature Learning plays a vital role in the implementation of ML techniques. These techniques are most appropriate and efficient for intrusion prevention and detection in the field of security [2][3][5]. Intrusion detection systems involving these techniques are based on the knowledge representation of the characteristics of network traffic. Most of the available benchmark datasets are high in dimension and imbalanced i.e. proportion of data from each class label is not equal, which may be affecting the performance of the classifier [6][7][8]. The anomalies in the network may be classified as high false positive [9] and may not be able to learn minor class labels.

Recently unsupervised DL models were widely used because of their affective representation of data for dimension reduction. Before using shallow ML techniques, an unsupervised Autoencoder model was used to extract optimal subspace to improve the detection rate and computational cost [10]. Passing too many features to ML algorithms would even lead to overfitting [11]. Several authors have proposed various NIDS in the literature using SAE for dimension reduction.

### 1.1. Problem Statement

Most of the authors have not concentrated on the following issues while using SAE.

- The problem of overfitting is not properly addressed.
- They failed to select a suitable activation function for SAE.
- They are not implementing suitable methodologies to identify the optimal hyperparameters.

To address the above issues the following techniques are considered in this paper.

- By imposing constraints on weights and dropout regularization to overcome the overfitting problem.
- To identify the best activation function by imposing the above two constraints.
- To identify the optimal hyperparameters through grid-based search technique.

To fulfill these research gaps an attempt has been made in this paper by combining Constraints Optimized Stacked Autoencoder (COSAE) for dimension reduction and Grid search based SVM-RBF classifier for evaluation of NIDS in Cloud environment (Figure 1). For conducting experiments CIICDS2017 dataset adopted because it contains contemporary attack types with a huge number of network flows in the Cloud environment.

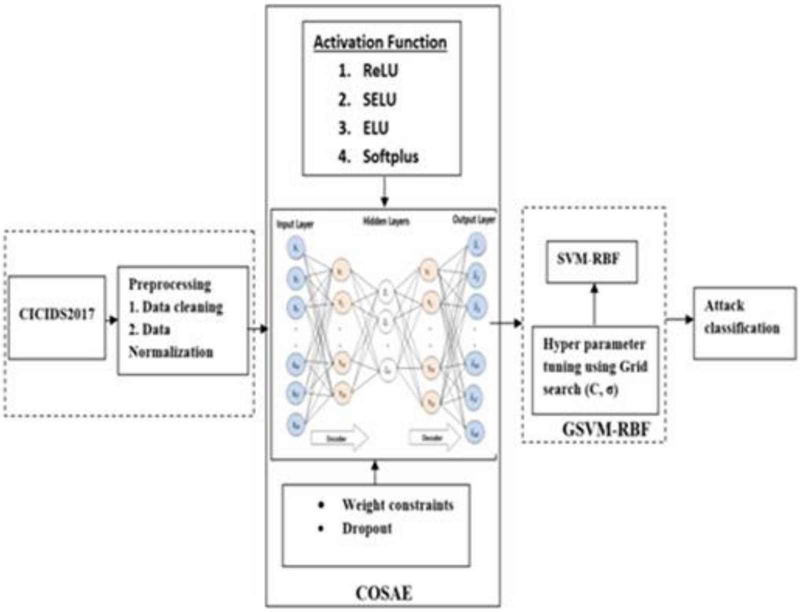

Figure 1 Proposed Framework of Integrating COSAE with GSVM-RBF Classifier

The motivation of the proposed work is to address the problem of overfitting and to select the suitable activation function while reducing the dimensionality for intrusion detection. Since the hyperparameters affect the performance of the model, the fine-tuning of of the SVM-RBF classifier is performed.

The remaining part of this paper is organized as follows. Section 2 presents a detailed description of the related research carried out in this area. Section 3 explains the Constraints Optimized Stacked Autoencoder (COSAE) for dimension reduction and Grid-search based SVM-RBF classifier for evaluation. Based on experiments, the results and discussions are presented in section 4, and finally in section 5 conclusions are drawn and the scope of the future work is discussed.

## 2. RELATED WORK

This section is intended to provide a briefing about the related work, contributions, and research gaps identified in them. A novel asymmetric Autoencoder (AE) based model was suggested by [12] to generate a latent space of dataset by virtue of minimizing the dimensionality and computational overhead. The main goal of this approach is to address the limited resources challenge and to lower the needed bandwidth to transfer data and computational overhead of low-end IoT devices. This auto-encoder is created with dense layers, followed by two LSTM layers and an MLP with 6 hidden layers. To minimize the computational overhead on the devices which are modeled as IoT a compression procedure is performed by applying distributive mode. The model is evaluated with three benchmark datasets namely NSL-KDD, UNSW-NB15, and TON IoT (UNSW-IoT20). The authors in this study didn’t focus on minimizing the loss while reconstructing the data.

The work presented in [13] detect DDoS attack in network traffic by using Autoencoder for dimension reduction and SVM-RBF classifier. The effectiveness of AE-SVM was also compared with the performance of SVM and PCA-SVM. Based on experimental results conducted on two datasets NSL-KDD and CICIDS2017 and they concluded that the proposed model yields better results when compared to other shallow ML techniques in terms of false-positive DDOS detection rate and rapid anomaly discovery which were high. The limitation of the proposed methodology is that it is tested with only one activation function.

A novel mathematical model based, practical intrusion detection application was proposed by [14] for practical intrusion detection. Their main intention was to tune the hyperparameters initially followed by the usage of Random Forest classifier to evaluate the model. Finally concluded that the results of the proposed IDS are more effective in terms of accuracy and F-measure when compared to recently published architectures for IDS.

The authors in [6] used various techniques for handling imbalanced dataset CIDDS-001. Initially, the synthetic minority reconstruction technique (SMRT) using variational autoencoder (VAE) as a classifier was employed to classify the original dataset. Later different techniques were employed for balancing class labels such as down sampling of majority class and up sampling of minor class. Various ML algorithms, deep neural networks, random forest, voting, variational autoencoder, and stacking classifiers were used against the resulting dataset to test the effectiveness of balanced data and assess the progress made in the direction of detecting the attacks. However, this study did not consider the computational overhead.

Whale optimization sparse autoencoder model (WO-SAE) based hyperparameter tuning of sparse autoencoder was proposed by [15] to improve the efficiency and performance of detection approach by extracting the low dimensional features. The various hyperparameters of SAE are sparsity weight, sparsity proportion, and the number of hidden layers, batch size, epochs, loss function, and optimizer. Along with Gated Recurrent Unit, Random Forest, DNN, and LSTM methods were applied to classify the time series data. They have concluded that their approach with Gated Recurrent Unit reduces the training time and minimizes the complexity of intrusion detection. However, the proposed methodology is tested on obsolete datasets which do not contain contemporary attacks of the cloud environment.

Denoising autoencoder with dropout regularization technique to the hidden layer was adopted for anomaly detection by [10] and experiments conducted on dataset NSL-KDD. They added noise data randomly not to copy the original input to output data. The input layer with 122 neurons followed by a dropout layer with 0.5 thresholds is applied. They have used a single layer structure for compression with a different number of neurons (8, 16, 24, and 32) and ReLU as activation function. Based on the results they proved that the intrusion detection rate is better than PCA and KPCA methods suggested by [16].

A Statistical Analysis Driven Optimized Deep Learning System is proposed for Intrusion Detection by combining two machine learning algorithms by [17]. In this paper, they adopted one hot encoding technique for transformation and subsequent cleaning so that it can be used and reduced to 50 features through AE with a scaled conjugate gradient optimizer. Then the reduced data is further classified using MLP with softmax layer. The proposed method achieves 87% of accuracy and proved that the accuracy is higher when compared to state-of-the-art models.

AE-RF based model is proposed for identifying the malicious traffic in [18]. The AE reduces high-dimensional data and extracts effective features. By testing various AE structures based on performance metrics and building time. The optimal structure is considered with a learning rate of 0.001,100 epochs and batch sizeis154 to reduce the dimension of CICIDS2017 from 78 to 16 features. Random Forest (RF) classifier is adopted to test the effect of dimension reduction. The accuracy of AE-RF is improved by 0.55% when compared to Pure RF.

The authors in [19] compared the performance of the shallow neural network (SNN) model and DNN model for cyber-attack detection. Before classifying the network traffic three different SAE models were implemented to reduce the features from 80 to 10, 20, and 30 of the CICIDS2017 dataset and evaluated with a softmax layer for multi-class classification. The 30 features yield the highest accuracies 96.71% for both AE-SNN and AE-DNN. This study did not pay attention to computational complexity.

Mighan, S. N., & Kahani, M. combined deep learning for dimension reduction and SVM classifier in [20] for improving the attack detection rate. In the first step, a deep autoencoder is utilized to reduce the features of the ISCX IDS UNB dataset from 42 to 10. In the next, step SVM-RBF is employed to construct the hyperplane to classify the attacks. The experimental results exhibit that the proposed DL-SVM outperforms with 90.2% accuracy when compared to the state-of-the-art models. This study does not consider the computational time as one of the performance metrics when compared with other models. This is identified as one of the important criteria for building the optimum model, so in this study to consider the computational time as one of the performance metrics.

The same authors in [21] proposed a hybrid SAE-SVM model for a fast and efficient intrusion detection system. Based on the different types of experimental results they concluded that the model is to improve the accuracy and also prediction speed when compared to the existing state-of-the-art models.

The work presented in [22] addresses the overfitting problem of neural networks on supervised learning tasks like vision, document classification, speech recognition, and computational biology on many benchmark datasets. Finally, they suggested that using dropout along with imposing max-norm weight constraints can optimize the performance of the neural network and prevent overfitting instead of just using dropout.

The present study is an extension of this study and addresses the overfitting problem on intrusion detection of the cloud environment. To achieve this objective with the implementation of regularization constraints such as weight constraints and dropout mechanisms. Several authors proposed different types of NIDS models with a combination of SAE for dimension reduction and CNN or DNN for classification. The overfitting problem occurs due to large neural nets and learning time is also becomes high.

By adopting two types of weight constraints orthogonality and unit norm to resolve the vanishing and exploding gradients. To avoid the overfitting problem the dropout layer is also implemented. While imposing these two constraints with dropout, the optimal model aims to improve the reconstruction error. Not much work is carried out to address these regularization techniques when building the NIDS models with a combination of SAE and various classifiers.

## 3. METHODOLOGY

In this section Constraints Optimized based Autoencoder (COSAE) is proposed to study the effectiveness of four activation functions for dimensionality reduction. The activation functions (ReLU, SELU, softplus, and ELU) are identified because of their prominent usage. The GSVM-RBF classifier is an enhancement to the existing SVM-RBF with the adoption of grid search for tuning of hyperparameters. This enhanced classifier is used to evaluate the COSAE through experimental results on the CICIDS2017 dataset. The proposed methodology consists of three phases namely i) Preprocessing, ii) Implementation of COSAE, and iii) GSVM-RBF classifier for evaluation.

### 3.1. Preprocessing

On the lines of [29], preprocessing techniques are adopted in this work which includes data cleaning, transformation, and normalization techniques, on the CICIDS2017 Wednesday dataset.

### 3.2. Implementation of COSAE

The COSAE designing process aimed at feature subset extraction includes the regularization techniques like weight constraints and dropout is as follows.

### 3.2.1. Weight Constraints Techniques

The training of deep neural networks is challenging by addressing overfitting problems [24], exploding gradients, and vanishing gradients [25]. The Weight constraints approach addresses these challenges and help to optimize model performance [25]. After going through these studies, the two weight constraints unit norm and orthogonality techniques are adopted to overcome exploding gradients, and vanishing gradients problem, and the same is explained below briefly.

Unit norm: This soft constraint is applied on both encoder and decoder; hence the weights are not exactly one, but they are adjusted near to one to avoid exploding gradients. Due to large values of weights the DNN loses its stability and to overcome this problem, this mechanism is adopted [25] which is as follows in Equation 1.

$$\sum_{i=1}^{k} W_{i,j}^{2} = 1 \text{Where i=1. . .. k} \quad (1)$$

Orthogonal weights: The weight orthogonality constraint on stacked autoencoder ensures that redundancy information is eliminated in encoder features and leads to a compact version of the network. The weight vectors are independent of each other and only informative weights are non-zero. Thus, only sufficient information flows through these non-zero weights during back-propagation and therefore vanishing gradients are avoided [11]. The equation of weight orthogonality constraint is enforced and the same is shown in Equation 2.

$$\mathrm{W^T\,W} = 1 \quad (2)$$

### 3.2.2. Dropout

While dealing with a large number of features and also with an imbalanced dataset DNN gets into an overfitting problem, due to this the learning time also increases. This problem is addressed by various authors through the dropout technique to minimize the regularization error. The adopted dropout technique is explained below on the lines of [26].

The introduction of the dropout regularizer in the autoencoder layer is aimed at increasing the generalization performance of the model. The key point of dropout is that it simply drops out units in network layers during training with specific probability p. By dropping a unit out means temporarily removing it from the network, along with all its incoming and outgoing connections. The choice of which units to drop is random. The dropout layer prevents the autoencoder from copying input to create output [27]. Based on the experimental results presented in [28] the typical possible values of dropout are in the range of 0 to 1. Hence in this study, the dropout layer is also added to SAE with a rate of 0.25 to combat the problem of overfitting. The optimized structure of SAE for dimension reduction identified in [29] is adopted to apply the aforementioned regularization constraints.

### 3.3. GSVM-RBF Classifier

This section discusses the GSVM-RBF classifier for multi-class classification.

The following Algorithm1 explains the process of finding the optimal values of c and σ using the grid-search technique. The input to Algorithm1 is the output of COSAE i.e. reduced dataset. The final leg of the algorithm is presented as Step 4 in which , the model is built on a train set and evaluated on a test set for each combination of C and σ. Accuracy is considered for each combination as a basic performance metric among them, highest accuracy is identified, then the corresponding values of c and σ are considered as optimal values.

The efficiency of the SVM classifier depends on i) which kernel has to be selected and ii) whether the values of the hyperparameters are optimized or not. The first one is fulfilled with the adoption of the RBF kernel because based on the experimental analysis conducted by [23]. To address the problem of hyperparameters tuning, c and σ are optimized with the adoption of Grid-based search algorithm [30].

The GSVM-RBF classification process is divided into two phases i) finding the optimal values for c and σ and ii) implementing multi-class classification. The c and σ input values are chosen in the following algorithm as per the earlier studies of [31] [4].

```
Input: D //dataset containing input features X and class label
         y post reduction of dimension, data sample
         X∈R where {x1, x2…., x30} represent different
         features.
         Output: Optimized values for c and σ
Step1: //initialization
                 C= {0.1, 1, 10, 100, 101}, //Set
                      representing the values of C
                 σ= {1, 0.1, 0.08, 0.09, 0.01, 0.001}, Set
                      representing the values of σ
                 K=5,     //Number of cross validations
                 best_accuracy=0, c_opt=0, σ_opt =0. // Assign
                 the initial values to variables
   Step2: Divide the data into K folds with approximately
                 equal distribution of labels
   Step3: for fold k_i in the K folds
                 set fold k_i as X_test, y_test set and
                 remaining K-1 folds as X_train and y_train
                 set
   Step4: for all c in the set C
                 for all σ in the set σ do
                      model=svm_train (X_train and
                           y_train, c, σ)
                    y_val=model.predict (X_test)
                    accuracy=svm.score (y_test, y_val)
                      //train svm for every pair of hyper
                      parameter
                 if accuracy>best_accuracy
                   best_accuracy=accuracy
                   c_opt =c
```

```
                    $\sigma_{opt} = \sigma$
                end if
            end for
        end for
    end for
Step 5: return $c_{opt}$, $\sigma_{opt}$
```

Algorithm1: Grid-based search algorithm to find the optimal values of c and σ

There are two techniques to deal with multi-class classification in SVM and they are one-versus-one” (OVO) and “one-versus-all” (OVA). As per [32] the OVO or pair wise classification of the CICIDS2017 dataset with six classes require fifteen binary models. It is more computationally intensive, whereas OVA approach takes only six models to distinguish all different classes [32-33]. Due to this reason the OVA approach is considered for conducting experiments with optimized values of $c_{opt}$, $\sigma_{opt}$ classify the attacks using GSVM-RBF kernel.

## 4. EXPERIMENTAL SETUP

The proposed methodology is implemented as a program using scikit-learn library of Python. Experiments are conducted to investigate the efficiency of COSAE in Google colab-hosted virtual machine, GPU Tesla P100-PCIE-16GB using TensorFlow as backend with Keras as a higher-level framework on Windows10 64-bit system. The GSVM-RBF multi-class classification is implemented on Intel core i5 1.80 GHz processor, 8GB RAM for validating with five-fold cross-validation for different c and σ values. The parameter settings of the COSAE model are given in Table 1.

| Model Parameters | Values |
|---|---|
| Input Neurons | 68 |
| Output Neurons | 68 |
| Number of hidden layers | 3 |
| Number of Neurons in hidden layer | 30 |
| Optimizer | Adadelta |
| Loss function | Mean Squared Error |
| Epochs | 10 |
| Batch size | 256 |
| Learning rate | 0.01 |
| Dropout | 0.25 |

Table 1 Hyperparameters of COSAE Model

In the similar lines of [29], the five layers 68-50-30-50-68 neurons are considered for implementing COSAE with the above regularization techniques. As per the earlier studies of [29][36] the experiments were also conducted on COSAE with four activation functions ReLU, SELU, Softmax, and ELU which were short-listed as optimized functions for further study.

In this experiment, the range of hyperparameter values used to tune are C and σ, which are represented as Sets. Five-fold cross-validation is applied with grid search, where four parts of data were used as training and the residual part was used to validate. The experiments were repeated five times similarly with different parts and finally, the parameters which result with the highest classification accuracy are returned as optimal parameters ($c_{opt}$=100, $\sigma_{opt}$=1).

## 5. RESULTS AND DISCUSSION

The problem of overfitting is not properly addressed by the earlier researchers. Failure to select a suitable activation function for SAE, and suitable methodologies are not done correctly that fits to identify the optimal hyperparameters. In order to overcome these issues, the proposed work addresses by imposing constraints on weights and dropout regularization to avoid the overfitting problem. The activation function is selected by imposing the above two constraints and the optimal hyperparameters identified through grid-based search technique. To evaluate the proposed hybrid model with the integration of COSAE and GSVM-RBF, various performance metrics namely Precision, F-Measure, Accuracy and, Computational time are computed over the CICIDS2017 dataset. The computational experiments are carried out in two scenarios based on with and without regularization constraints using four activation functions and compared with SVM-RBF and GSVM-RBF classifiers. When comparing both the scenarios the performance of SAE and COSAE are evaluated. By the experimental results, scenario2 COSAE+GSVM-RBF yields satisfactory results.

### Scenario 1

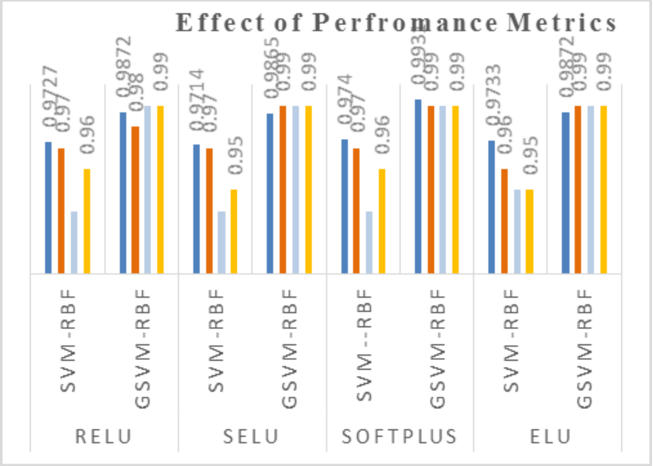

Figure 2 Influence of Various Performance Metrics for Different Activation Functions of Secnario1

The experiments are conducted in secnario1 with SAE+SVM-RBF and SAE+GSVM-RBF hybrid classification models and obtained performance metrics with computational time are shown in Tables 2 and 3 and Figures from 2 to 5.

From Figure 2 and Table 2 it is observed that the performance metrics and gain in computational time give the highest values with GSVM-RBF when compared to SVM-RBF. Among all the combinations tested, the softplus activation function is the best performer in terms of accuracy for both models and ELU is the nearest second.

| Activation Function | Model | Accuracy | Precision | Recall | F-Measure | Computational Time (Sec.) |
|---|---|---|---|---|---|---|
| ReLU | SVM-RBF | 0.9727 | 0.97 | 0.94 | 0.96 | 349.6 |
| | GSVM-RBF | 0.9872 | 0.98 | 0.99 | 0.99 | 144.31 |
| SELU | SVM-RBF | 0.9714 | 0.97 | 0.94 | 0.95 | 333.33 |
| | GSVM-RBF | 0.9865 | 0.99 | 0.99 | 0.99 | 158.41 |
| Softplus | SVM-RBF | 0.974 | 0.97 | 0.94 | 0.96 | 338.25 |
| | GSVM-RBF | 0.9931 | 0.99 | 0.99 | 0.99 | 141.63 |
| ELU | SVM-RBF | 0.9733 | 0.96 | 0.95 | 0.95 | 306.59 |
| | GSVM-RBF | 0.9872 | 0.99 | 0.99 | 0.99 | 149.5 |

Table 2 Influence of Performance Metrics and Computational Time for Different Activation Functions of Scenario1

All the activation functions tested exhibit the performance in similar lines with a value of 0.99 for precision, recall, and F-Measure of GSVM-RBF. Whereas SVM-RBF exhibits differently across different activation functions, with precision, and recall of the ELU has a difference of 0.01. The other activation functions are equally performing with the values of 0.97 and 0.94. There is an insignificant difference is identified between these activation functions for both classification models.

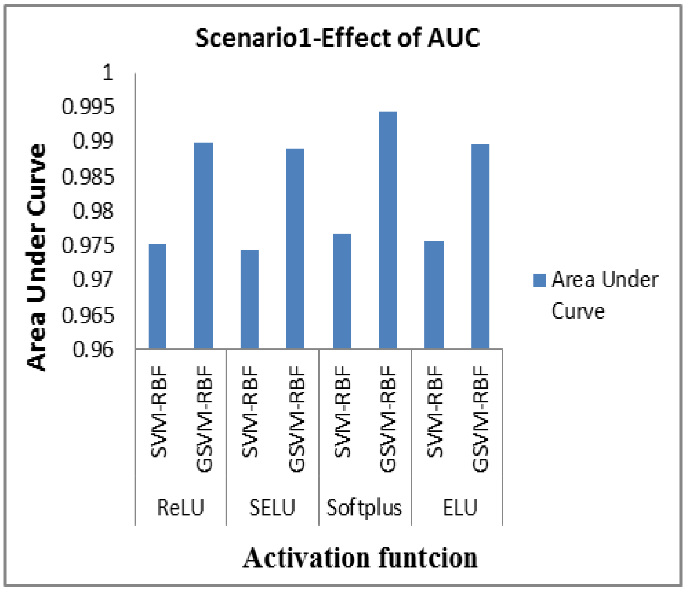

Figure 3 Influence of AUC for Different Activation Functions of Scenario1

It can be observed from Figure 3 and Table 3 that the GSVM-RBF classifier gives better values of AUC for all activation functions when compared to SVM-RBF. The gain in AUC for softplus is the highest value with 1.8%.

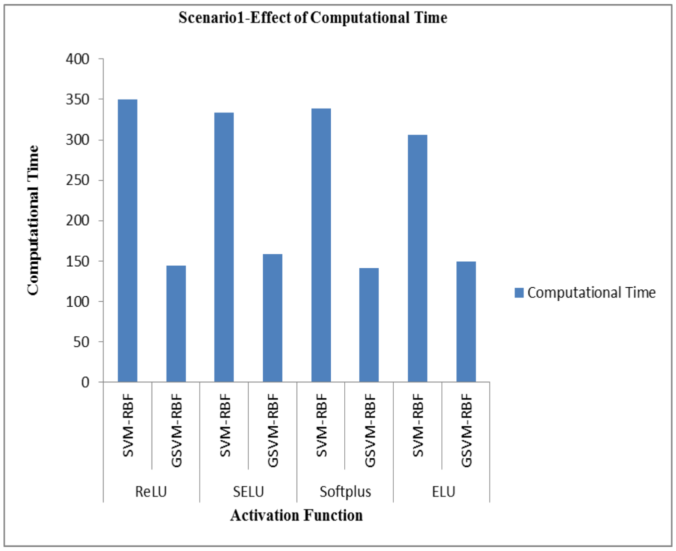

Figure 4 Influence of Computational Time for Different Activation Functions of Secnario1

The above Figure 4 shows that the gain in computational time gives the highest values with GSVM-RBF when compared to SVM-RBF. The gain in computational time for GSVM-RBF is better when compared to SVM-RBF for all activation functions with softplus at the first place and ReLU in the second place.

| Activation Function | Model | Area Under Curve | AUC gain in percentage with GSVM-RBF |
|---|---|---|---|
| ReLU | SVM-RBF | 0.9752 | 1.5 |
| | GSVM-RBF | 0.9898 | |
| SELU | SVM-RBF | 0.9744 | 1.5 |
| | GSVM-RBF | 0.9891 | |
| Softplus | SVM-RBF | 0.9767 | 1.8 |
| | GSVM-RBF | 0.9945 | |
| ELU | SVM-RBF | 0.9758 | 1.4 |
| | GSVM-RBF | 0.9897 | |

Table 3 Influence of AUC for Different Activation Functions of Secnario1

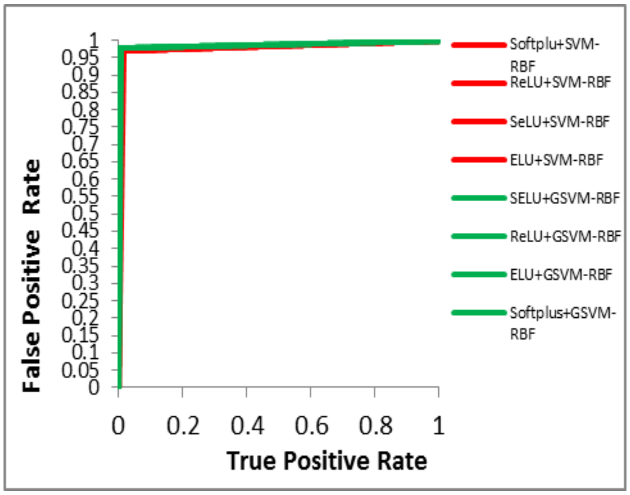

Figure 5 Influence of ROC for Different Activation Functions of Scenario1

A steeper curve towards the y-axis of Figure 5 indicates that the SAE+GSVM-RBF has better intrusion detection capability for multi-class classification compared to SAE+SVM-RBF. The behavior of various activation functions regarding ROC follows the same fashion and is close to each other.

Based on the above observations it can be concluded that the softplus is a better activation function for SAE w.r.t dimension reduction with minimum computational time and without compromising on accuracy through GSVM-RBF classifier for ideal NIDS.

**Scenario2**

This scenario aims to investigate the effect of COSAE with the combinations of SVM-RBF and GSVM-RBF classifiers. The experimental results are presented in tables 4 and 5 and the relevant graphs are given from figures 6 to 9.

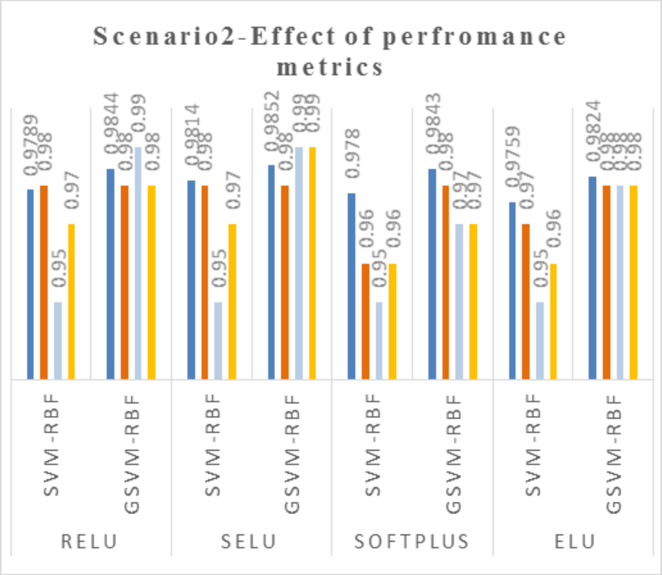

Figure 6 Influence of Performance Metrics for Different Activation Functions of Secnario2

The following observations are drawn from Figure 6 and Table 4. The performance metrics and computational time of GSVM-RBF at the first glance give an impression of better values when compared to SVM-RBF for all activation functions. For both classifiers, the SELU performs with better accuracy when compared to other activation functions. The ReLU and softplus are in the next place with more or less equal performance. The ReLU and SELU perform equally and provide better values for precision, recall, and F-Measures.

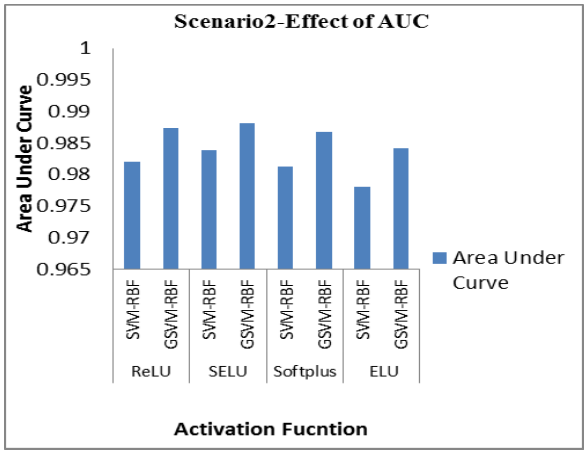

Figure 7 Influence of AUC for Different Activation Functions of Secnario2

| Activation Function | Model | Accuracy | Precision | Recall | F-Measure | Computational Time (Sec.) |
|---|---|---|---|---|---|---|
| ReLU | SVM-RBF | 0.9789 | 0.98 | 0.95 | 0.97 | 202.36 |
| | GSVM-RBF | 0.9844 | 0.98 | 0.99 | 0.98 | 187.03 |
| SELU | SVM-RBF | 0.9814 | 0.98 | 0.95 | 0.97 | 178.79 |
| | GSVM-RBF | 0.9852 | 0.98 | 0.99 | 0.99 | 121.5 |
| Softplus | SVM-RBF | 0.978 | 0.96 | 0.95 | 0.96 | 233.08 |
| | GSVM-RBF | 0.9843 | 0.98 | 0.97 | 0.97 | 210.48 |
| ELU | SVM-RBF | 0.9759 | 0.97 | 0.95 | 0.96 | 204.34 |
| | GSVM-RBF | 0.9824 | 0.98 | 0.98 | 0.98 | 112.93 |

Table 4 Influence of Performance Metrics and Computational Time for Different Activation Functions

| Activation Function | Model | AUC | Percentage gain using GSVM-RBF |
|---|---|---|---|
| ReLU | SVM-RBF | 0.9821 | 0.5 |
| | GSVM-RBF | 0.9874 | |
| SELU | SVM-RBF | 0.9838 | 0.6 |
| | GSVM-RBF | 0.9881 | |
| Softplus | SVM-RBF | 0.9813 | 0.4 |
| | GSVM-RBF | 0.9868 | |
| ELU | SVM-RBF | 0.9781 | 0.4 |
| | GSVM-RBF | 0.9841 | |

Table 5 Influence of AUC for Different Activation Functions of Secnario2

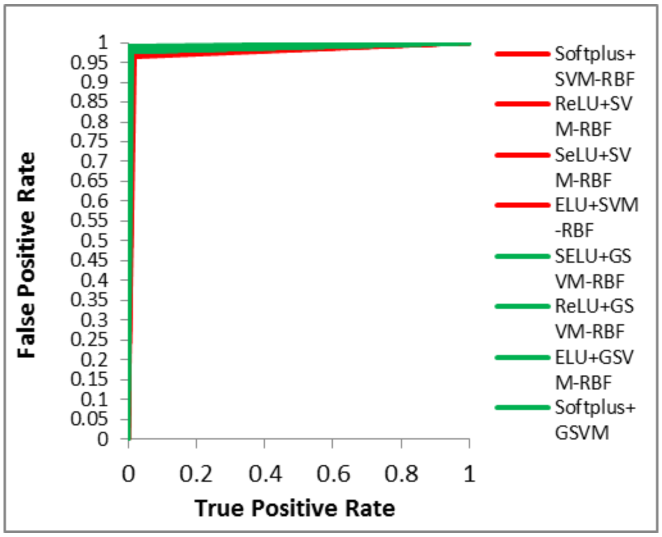

Figure 8 Influence of ROC for Different Activation Functions of Scenario2

From Table 5 and figures 7 and 8, it is observed that the SAE+GSVM-RBF performs well when compared to SAE+SVM-RBF for all activation functions. The SELU achieves the highest gain of AUC with a value of 0.6%. The implementation of regularization constraints enhances the performance of detection rate in minority classes. The generated ROC result shows that the COSAE+GSVM-RBF is the best classification model with a high true positive rate. The ROC curves for both the methods for all activation functions follow a similar kind of performance.

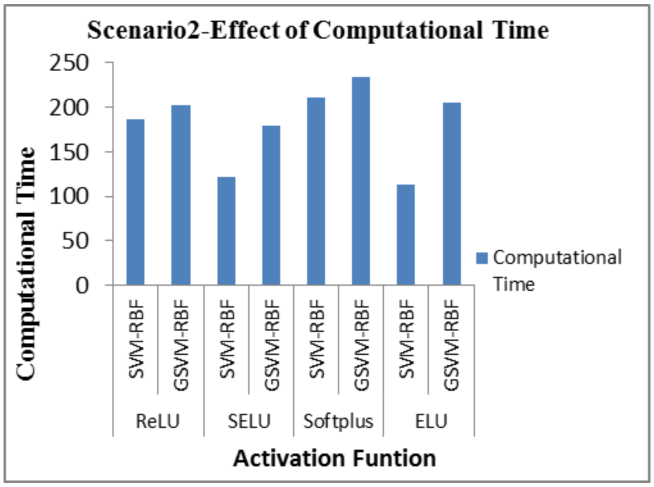

Figure 9 Influence of Computational Time for Different Activation Functions of Secnario2

The computational time of GSVM-RBF shown in Figure 9 illustrates that it is minimum when compared to SVM-RBF for all activation functions. The overhead in computational time gets reduced with ELU for GSVM-RBF and the next lowest is SELU. Based on the above observations it can be concluded that COSAE+GSVM-RBF is a better classifying model for building an ideal NIDS. The SELU has a better standing in case of the performance metrics. ELU shows a considerable improvement of gain in the context of computational time, when COSAE is used for dimension reduction.

## 6. COMPARISION WITH OTHER STATE-OF-THE-ART MEHTODS

The evaluated performance metrics of the proposed COSAE+GSVM-RBF model are compared with the other

State-of-the-art intrusion models based on conventional ML and deep learning techniques. Table 6 shows various performance metrics that are evaluated using the CICIDS2017 dataset. The comparison is carried out based on four classification metrics accuracy, precision, recall, and F-measure.

In [19] Ustebay, S. et al. reduced the CICIDS dataset to 10,20, and 30 features respectively with ReLU activation function using three different AE models. Further, the reduced data was trained with SNN and DNN classifiers on each dataset. Among these results, the highest accuracies were obtained with 30 features i.e., 95.12 and 96.71 only. Kumar, V. et al. in [34] applied various classification algorithms to detect DDoS attacks in the CICIDS-2017 dataset and were successive with ID3 classifier as 95%. In [35] authors enhanced the performance of the Ada-boost classifier by addressing the imbalance of the dataset with an ensemble feature selection method. The achieved results are with accuracy 81.83%, precision of 81.83%, recall of 100%, and F1-Score of 90.01%.

By observing the above Table 6, the COSAE+GSVM-RBF model is a better model when compared with other state-of-the-art methods performances by 98.52% accuracy. The deep feature learning methods yield better values when compared to those without feature learning. It is a clear manifestation that feature learning through COSAE is an important pre-processing task for applying ML/DL algorithms along with tuning of hyperparameters for NIDS to achieve better results.

| Study | Type | Method | Accuracy | Precision | Recall | F-Measure |
|---|---|---|---|---|---|---|
| [34] | Conventional ML methods | Naïve Bayes | 0.82 | 0.82 | 0.8165 | 0.82 |
| | | ID3 | 0.95 | 0.95 | 0.9515 | 0.95 |
| | | Ensemble Learning | 0.89 | 0.89 | 0.891 | 0.89 |
| | | MLP | 0.84 | 0.84 | 0.8387 | 0.84 |
| [35] | Feature Learning + Conventional ML method | AdaBoost + EFS + SMOTE | 0.8183 | 0.8183 | 1 | 0.90 |
| [19] | Feature Learning and classification based on Deep Learning | AE+SNN | 0.9512 | - | - | - |
| | | AE+DNN | 0.9671 | - | - | - |
| Proposed Approach | Feature Learning with Deep Learning and conventional classification method | COSAE+GSVM-RBF | 0.9852 | 0.98 | 0.99 | 0.99 |

Table 6 The Comparison with State-of-the-Art Methods

## 7. CONCLUSION

The current research work proposes a framework for NIDS in the cloud environment with the integration of deep feature learning and conventional classification techniques. COSAE is used as a deep feature learning technique with weight and dropout regularization constraints. Four activation functions ReLU, SeLU, softplus, and ELU are considered for feature learning with COSAE to arrive at the best activation function. The classifier is different from traditional SVM because the hyperparameters are optimized based on grid search. For validation of this model, the CICIDS2017 dataset is adopted for conducting experiments. It is observed that the gain of computational time of the SAE+GSVM-RBF model is approximately two times of the SAE+SVM-RBF model with marginal differences of the all-performance measures. This is true for all activation functions among them softplus gives an additional gain in computational time. To observe that the gain in computational time of the COSAE+GSVM-RBF model is higher than the COSAE+SVM-RBF model with insignificant differences of all performance measures. Among the activation functions, the gain of the computational time with ELU is maximum and without compromising on the accuracy front. The results of AUC for SAE+GSVM models are high for all activation functions and the softplus exhibits the highest gain among them. It is observed that the gain of AUC using SeLU is high using the COSAE+GSVM model. Finally, it is concluded that the integrated COSAE+GSVM-RBF model is suitable to build an ideal NIDS to detect the attacks with minimum computational time and the accuracy of the model is also good when compared to other models. In summary, the COSAE+GSVM works well with an imbalanced dataset and is strong enough to detect attacks. As a future scope of the research, this model may be compared with other kernel and activation functions with various hyperparameter optimization techniques and the evolutionary process may be conducted on current benchmark datasets containing a new type of malicious activities. This model may be evaluated on real-time traffic testbeds. It can be observed that in the COSAE+GSVM-RBF model with ELU activation function, the prediction time is minimized when compared to other combinations of SAE, COSAE, SVM-RBF, and GSVM-RBF with four activation functions. This study concludes that COSAE+GSVM-RBF with ELU activation function is better for building NIDS with a minimum prediction time and non-significant difference of other performance measures when compared to combinations of SAE, COSAE, SVM, and GSVM. As a future enhancement, this study may be evaluated on real-time network traffic.

## REFERENCES

[1] Shawahna, A., Abu-Amara, M., Mahmoud, A. S., & Osais, Y. (2018). “EDoS-ADS: an enhanced mitigation technique against economic denial of sustainability (EDoS) attacks”. IEEE Transactions on Cloud Computing, 8(3), 790-804.

[2] Tulasi Bhavani, T., Rao, M.K., Reddy, A.M. (2020). “Network intrusion detection system using random forest and decision tree machine learning techniques”. Advances in Intelligent Systems and Computing, Issue-1045. pp. 637-643.

[3] Krishna Anne, V.P., Rajasekhara Rao, K. (2017). “Standards and analysis of intrusion detection-based system: A comparative study”. ponte, Volume-73 Issue-2, pp. 87-97.

[4] Aamir, M., & Zaidi, S. M. A. (2019). “Clustering based semi-supervised machine learning for DDoS attack classification”. Journal of King Saud University-Computer and Information Sciences.

[5] Jadhav, A.D., Pellakuri, V. (2019). “Performance analysis of machine learning techniques for intrusion detection system”." International conf. on Computing, Communication Control and Automation, ICCUBEA 2019.

[6] Abdulhammed, R., Faezipour, M., Abuzneid, A., & AbuMallouh, A. (2018). “Deep and machine learning approaches for anomaly-based intrusion detection of imbalanced network traffic”. IEEE sensors letters, 3(1), 1-4.

[7] Zhang, Y. (2018). “Deep generative model for multi-class imbalanced learning”. M.S. thesis,Dept. Elect., Comput., Biomed. Eng., Univ. Rhode Island, South Kingston, RI, USA,2018.

[8] Chawla, N. V. (2009). “Data mining for imbalanced datasets: An overview”. In Data mining and knowledge discovery handbook (pp. 875-886). Springer, Boston, MA.

[9] Karatas, G., Demir, O., & Sahingoz, O. K. (2020). “Increasing the Performance of Machine Learning-Based IDSs on an Imbalanced and Up-to-Date Dataset”. IEEE Access, 8, 32150-32162.

[10] Mohamed, S., Ejbali, R., & Zaied, M. (2019). “Denoising Autoencoder with Dropout based Network Anomaly Detection”. ICSEA 2019, 110.

[11] Bhardwaj, A., Mangat, V., & Vig, R. (2020). “Hyperband Tuned Deep Neural Network with Well Posed Stacked Sparse AutoEncoder for Detection of DDoS Attacks in Cloud”. IEEE Access, 8, 181916-181929.

[12] Andalib, A., & Vakili, V. T. (2020).” A Novel Dimension Reduction Scheme for Intrusion Detection Systems in IoT Environments”. arXiv preprint arXiv:2007.05922.

[13] KASIM, Ö. (2020).” An efficient and robust deep learning based network anomaly detection against distributed denial of service attacks”. Computer Networks, 180, 107390.

[14] Musafer, H., Abuzneid, A., Faezipour, M., & Mahmood, A. (2020). “An Enhanced Design of Sparse Autoencoder for Latent Features Extraction Based on Trigonometric Simplexes for Network Intrusion Detection Systems”. Electronics, 9(2), 259.

[15] Xiaopeng, C., & Hongyan, Q. “Deep feature Extraction Via Sparse Autoencoder for Intrusion Detection System”,2020.

[16] Elkhadir, Z., Chougdali, K., & Benattou, M. (2016). “Intrusion detection system using pca and kernel pca methods”. In Proceedings of the Mediterranean Conf. on Information & Communication Technologies 2015 (pp. 489-497). Springer, Cham.

[17] Ieracitano, C., Adeel, A., Gogate, M., Dashtipour, K., Morabito, F. C., Larijani, H., ... & Hussain, A. (2018, July). “Statistical analysis driven optimized deep learning system for intrusion detection”. In International Conf. on Brain Inspired Cognitive Systems (pp. 759-769). Springer, Cham.

[18] Yeom, S., Choi, C., & Kim, K.” AutoEncoder Based Feature Extraction for Multi-Malicious Traffic Classification”.2020.

[19] Ustebay, S., Turgut, Z., & Aydin, M. A. (2019, June). “Cyber Attack Detection by Using Neural Network Approaches: Shallow Neural Network, Deep Neural Network and AutoEncoder”. In International Conf. on Computer Networks (pp. 144-155). Springer, Cham.

[20] Mighan, S. N., & Kahani, M. (2020). “A novel scalable intrusion detection system based on deep learning”. International Journal of Information Security, 1-17.

[21] Mighan, S. N., & Kahani, M. (2018, May). “Deep learning based latent feature extraction for intrusion detection”. In Electrical Engineering (ICEE), Iranian Conf. (pp. 1511-1516). IEEE.

[22] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). “Dropout: a simple way to prevent neural networks from overfitting”. The journal of machine learning research, 15(1), 1929-1958.

[23] Narisetty, N., Kancherla, G. R., Bobba, B., & Swathi, K.” Performance of Various SVM Kernels for Intrusion Detection of Cloud Environment”. IJETER, volume 8, No.10, October 2020.

[24] Vorontsov, E., Trabelsi, C., Kadoury, S., & Pal, C. (2017, July). “On orthogonality and learning recurrent networks with long term dependencies”. In International Conf. on Machine Learning (pp. 3570-3578). PMLR.

[25] Courtenay, L. A., Huguet, R., Gonzalez-Aguilera, D., & Yravedra, J. (2020). “A hybrid geometric morphometric deep learning approach for cut and trampling mark classification”. Applied Sciences, 10(1), 150.

[26] Livieris, I. E., Iliadis, L., & Pintelas, P. (2020). “On ensemble techniques of weight-constrained neural networks”. Evolving Systems, 1-13.

[27] Sadaf, K., & Sultana, J. (2020). “Intrusion Detection Based on Autoencoder and Isolation Forest in Fog Computing”. IEEE Access, 8, 167059-167068.

[28] Alexander Pauls & Josiah A Yoder (2018). “Determining Optimum Drop-out Rate for Neural Networks”. http://micsymposium.org/mics2018/proceedings/MICS_2018_paper_27.pdf.

[29] Narisetty, N., Kancherla, G. R., Bobba, B., & Swathi, K. “Investigative Study of the Effect of Various Activation Functions with Stacked Autoencoder for Dimension Reduction of NIDS using SVM”. IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 12, No. 5, 2021.

[30] Budiman, F. (2019). “SVM-RBF parameters testing optimization using cross validation and grid search to improve multiclass classification”. Научная визуализация, 11(1), 80-90.

[31] Biggio, B., Fumera, G., & Roli, F. (2010, April). “Multiple classifier systems under attack”. In International workshop on multiple classifier systems (pp. 74-83). Springer, Berlin, Heidelberg.

[32] Wang, W., Du, X., Shan, D., & Wang, N. (2019, October). “A Hybrid Cloud Intrusion Detection Method Based on SDAE and SVM”. In 2019 12th International Conf. on Intelligent Computation Technology and Automation (ICICTA) (pp. 271-274). IEEE.

[33] Hsu, C. W., & Lin, C. J. (2002).” A comparison of methods for multiclass support vector machines”. IEEE transactions on Neural Networks, 1 3(2), 415-425.

[34] Kumar, V., Choudhary, V., Sahrawat, V., & Kumar, V. (2020, June). “Detecting Intrusions and Attacks in the Network Traffic using Anomaly based Techniques”. In 2020 5th International Conf. on Communication and Electronics Systems (ICCES) (pp. 554-560). IEEE.

[35] Yulianto, A., Sukarno, P., & Suwastika, N. A. (2019, March). “Improving adaboost-based intrusion detection system (IDS) performance on CIC IDS 2017 dataset”. In Journal of Physics: Conf. Series (Vol. 1192, No. 1, p. 012018). IOP Publishing.

[36] Narisetty, N., Kancherla, G. R., Bobba, B., & Swathi, K (2021, April).” Performance Analysis of Different Activation and Loss Functions of Stacked Autoencoder for Dimension Reduction for NIDS on Cloud Environment”. International Journal of Engineering Trends and Technology Volume 69 Issue 4, 169-176, April 2021 ISSN: 2231 – 5381 /doi:10.14445/22315381/IJETT-V69I4P224

Authors

**N. Nirmalajyothi**, is a Ph. D. Research scholar in the Dept. of Computer Science and Engineering, Acharya Nagarjuna University, Guntur, India.She has received M.Tech (2010) degree in CSE from St.Marys College of Engineering & Technology,Hyderabad, and published about 10 research papers in international journals in the fields of Network Security, Cloud Computing Security, Machine learning, etc.

**Dr. K Gangadhara Rao,** received his Ph.D. in 2011, and about 6 Ph.D. scholars were awarded under his guidance and guiding about 8 Ph.D. scholars in the Dept. of Computer Science and Engineering, Acharya Nagarjuna University, Guntur, India. Currently working as professor & Head of Dept. of Computer Science and Engineering. His Interesting research areas are Cloud Computing, Manets, Software Engineering, Data Mining etc and published 37 papers in various national/international journals and Conferences.

**Dr. B. Basaweswara Rao**, received his Ph.D. in 2004, and about 7 Ph.D. scholars were awarded under his guidance and guiding about 8 Ph.D. scholars in the Dept. of Computer Science and Engineering, Acharya Nagarjuna University, Guntur, India. Currently working in areas like Network Security, Cloud Computing, Mobile Ad-hoc Networks, Big data Analytics, game theory, etc. and published 30+ papers in various national/international journals and Conferences.

**Dr. K. Swathi**, received her Ph.D. in 2019 in Computer Science and Engineering from Acharya Nagarjuna University, Guntur, India. Currently Professor in the Department of CSE, KL University, Guntur. She has published about ten international journals and about ten national/international conferences in the fields of Network Security, Data Analytics, Machine Learning, Cloud Security, Technology in Education, etc.

**How to cite this article:**

Nirmalajyothi Narisetty, Gangadhara Rao Kancherla, Basaveswararao Bobba, K. Swathi, “Hybrid Intrusion Detection Method Based on Constraints Optimized SAE and Grid Search Based SVM-RBF on Cloud”, International Journal of Computer Networks and Applications (IJCNA), 8(6), PP: 776-787, 2021, DOI: 10.22247/ijcna/2021/210725.