Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

Neural Network for Automatic Encryption Using Key Component Analysis to Detect Network Intrusion in Cloud Computing

Jenan jader msaad[1,*], Alaa majeed shnin[2]
[1]Al-Furat Al-Awsat Technical University / Kufa- Department of computer science
Jenan.jader@atu.edu.iq
[2]Al-Furat Al-Awsat Technical University / Kufa- Babylon Technical Institute
alaa.shnen.iba@atu.edu.iq

https://doi.org/10.46649/fjiece.v3.2.6a.15.5.2024

Abstract. *Cloud computing is one of the networks that has attracted more people's attention than other networks in today's world, and the reason for this is that it stores data in itself. This network consists of different computers that are scattered in different places and each user will be able to be a member in this space by registering in it. Among the many important issues related to the environment is the issue of security, which has become a very complex challenge. In order to secure networks, tools have been created, and one of these tools is the use of artificial intelligence methods to establish security in this environment. In most of the researches carried out in this field, either the duration of the algorithm execution was very long or the researches did not have enough accuracy, so article, we tried to detect the penetration in cloud computing. This was done in two different scenarios for feature selection and reduction. In one scenario, we reduced the dimensions using the fundamental analysis method, and in the other scenario, we did this using the bat meta-heuristic algorithm. In both scenarios, we used self-encrypting neural network for classification, and the structure of this network was the same in both scenarios. The reason for using these two scenarios was that we want to compare the accuracy obtained and the execution time of the algorithm for the dimensionality reduction method with principal component analysis and meta-heuristic algorithm. , the obtained results showed that the accuracy of the dimension reduction method with the meta-heuristic algorithm was more accurate than the principal component analysis method, but its execution time was approximately 13 minutes more than the principal component analysis method.*

Keywords: *security, cloud computing, principal component analysis, bat algorithm, self-encrypting neural network.*

1. INTRODUCTION

Cloud computing is a next-generation Internet-based system that will provide convenient and customized services to users to access or work with various cloud programs. Cloud computing will provide a way to record and access data from anywhere by connecting the cloud application to the Internet. By choosing cloud services, users can store local data in a remote data server. Information stored in the remote data center may be accessed or managed by cloud services provided by cloud service providers. Therefore, the processing of information stored in a remote server should be done with the utmost care. Cloud computing security is the main concern that is addressed today.

Although the issue of storing information on the virtual level will create a good ability and cost for users from the point of view of providing space to record information well, but they are still not able to fully satisfy. Most companies and organizations are either not familiar with this technology, or if they are

relatively familiar, the first thing that will appear in their minds is that this data is not safe from hacking. But it is interesting to know that cloud computing includes several basics and security issues, and they are sure that famous companies such as Google and Microsoft will provide two good services for this purpose. It took time to come to the conclusion that this service has been launched or not. Therefore, they will naturally reach an acceptable level of information security, which will now be used as an independent service for storage [1].

2- Related work

The rapid development of Internet technologies has dramatically increased the number of connected devices. This has created a large attack surface that requires the deployment of effective and practical countermeasures to protect network infrastructure from the damage that cyber attacks can cause. Hence, there is an absolute need to differentiate the boundaries in personal information and cloud and fog computing globally and to adopt specific information security policies and regulations. The purpose of the security policy and framework for cloud and fog computing is to protect end users and their information, reduce task-based operations, help enforce compliance, and establish standards for expected user actions, all of which are usage-based. Rules set for cloud computing In addition, intrusion detection systems are extensive solutions for monitoring and analyzing network traffic and detecting anomalies that can help identify ongoing hostile activities, provide alerts, and automatically block traffic from hostile sources.

In the article [2], it deals with the above issues by proposing an attention-based recurrent convolutional neural network (RCNN). This proposed RCNN is used to detect intrusive or non-intrusive text data. The plaintext information is then used for further processing and encrypted using a two-way encryption scheme. In this work, the elliptic curve encryption (ECC) approach is introduced to increase the performance of non-penetrative data security level.

In the paper [3], a new deep learning model based on convolutional neural networks and recurrent neural networks is developed for intrusion detection for cloud security. The proposed model was trained and tested using the NSL-KDD train dataset. With the proposed deep learning model, any detected and unverified traffic is prevented from reaching the cloud server.

In the study [4], artificial intelligence techniques, for example particle swarm optimization algorithm, MLP network are used to detect intrusions and attacks. The methods have been tested on NSL-KDD, KDD-CUP datasets.

In the article [5], an effective malware detection method in cloud infrastructure using Convolutional Neural Network (CNN) is discussed.

In [6], attackers are tracked using user interaction behavior pattern and deep learning technique. The actual user's mouse movements and clicks and keystrokes are stored in a database. Deep Belief Neural Network is designed using Restricted Boltzmann Machine (RBM) so that the RBM layer communicates with previous and subsequent layers.

3- Proposed method

, to increase the accuracy of intrusion detection systems in cloud environments, we propose to reduce the dimensions of the data in two different scenarios, in one scenario using the component analysis method Basically, we reduce the dimensions, and in the second scenario, we reduce the dimensions of the data using the bat evolutionary algorithm by selecting the appropriate feature, and then classify the pre-processed data using the self-encrypting deep neural network. Autoencoder network is an unsupervised network. The difference between such networks and other unsupervised networks is that the autoencoder network itself does not use the probability distribution for unsupervised training, but works in the same way as the supervised network. It means that it has a goal and is trained using gradient descent and error back propagation method. Autoencoder neural network is a functional network because it is a type of deep neural network that is used for feature extraction and reconstruction operations. This neural network has good performance and high accuracy. The purpose of adopting these two scenarios, is to compare statistical

methods such as principal component analysis with meta-heuristic methods for dimensionality reduction and to analyze the effect of these methods on the accuracy of the autoencoder classifier. , to better compare the results of both scenarios in terms of accuracy and algorithm execution time, both scenarios will be performed in the same conditions with the same network structure and also on the same database.

3-1- Scenarios of the proposed method

The purpose of adopting two scenarios , is to compare statistical methods such as principal component analysis with meta-heuristic methods for dimensionality reduction and to analyze the effect of these methods on the accuracy of the autoencoder classifier. , we will explain both scenarios.

3-1-1- The first scenario of the proposed method

In this scenario, we use the feature selection method using the bat algorithm to reduce the dimensions of the data. The steps of this scenario are divided into three steps after receiving the information, these three steps are: pre-processing, reducing the dimensions of features using feature selection by bat algorithm, classification using auto-encrypting neural network. , the structure of this scenario will be shown in the block diagram of the proposed method in Figure (1), then all the steps of the proposed method will be explained.

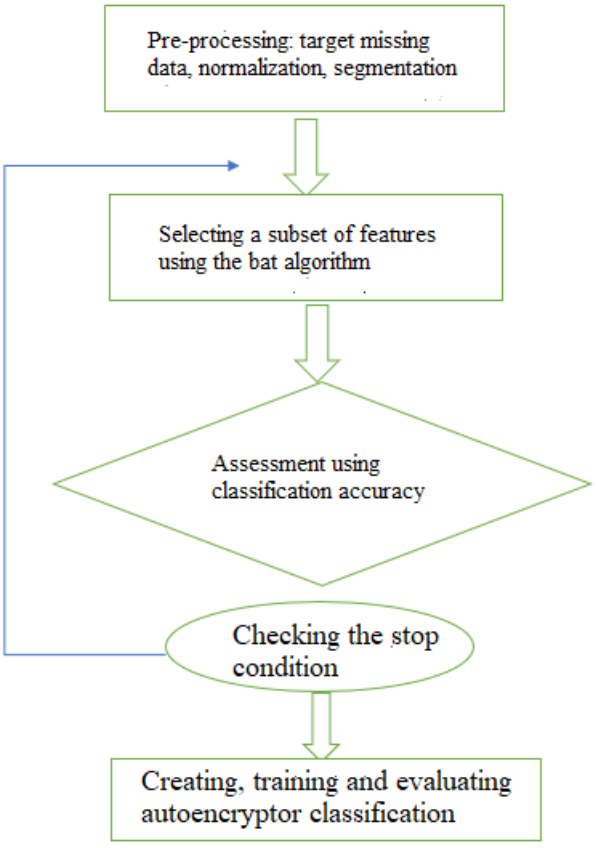

Figure 1: Steps of the proposed method of the first scenario

3-1-1-1- Data preprocessing

This stage includes several other sub-neighborhoods. which include removing missing data, normalizing data, separating the features in the dataset from labels and dividing the data into training and testing categories.

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

,, to process missing data, using averaging over the sum of the values of the neighbors of the missing data, we replace the obtained value as the missing data value. In the next step, the data are normalized. This step causes the information to be placed in a specific interval, which is used for this purpose from equation (1), in which X_i means the real value of the data and X is the normalized value:

$$X_n = \frac{X_i - X_{min}}{X_{max} - X_{min}}(H - L) + L,\ i = 1,2,\dots,N \quad (1)$$

3-1-1-2- Dimension reduction using the bat algorithm

The steps of implementing the bat algorithm are described as follows:

1- Creating a number of initial bats and initializing the bats

For all bats, we create a vector that is then equal to the number of features in the original database. Each dimension represents a feature. In the following text, in the first step of creating these vectors, they are given random values. We consider these values as the score of each feature. The higher the score of each feature, the higher its importance to be selected.

2- Evaluation of the fitting function for all bats

First, in this step, N features with a higher score are selected, and by using those features, we create a new data set. For the purpose of the new database, we evaluate the fit value for that bat by using the self-encrypting classifier.

3- Finding a bat that has the best fitting function

According to the classification, the subset that has the best accuracy has a better fitting function against the rest of the bats.

4- Movement of bats towards a bat with a better fitting function

Bats move towards a better-fitting bat in a situation where the loudness of the reflected sound is sufficiently loud. This work is combined by creating a random number so that the property of being random is considered in the algorithm and the bats are not trapped in the local optimum.

The value assigned to each bat dimension changes as bats move towards the bat closer to the prey.

5- If the frequency value has the default value, the work is finished, otherwise, it goes back to the second step.

6- Finding a bat with a better fit.

After the number of iterations of the algorithm is over, a bat is selected that has a better fitting function value. By applying the selected features, it creates a dataset bat that can be used instead of the original dataset.

3-1-1-3- Classification using autoencoding neural network

An autoencoder may have three or more layers: input, hidden, and output layers. In the simplest case, an autoencoder consists of an encoder and a decoder with only one hidden layer.

The input is given to the encoder and the output is extracted from the decoder. In this type of networks, instead of training the network and predicting the amount of the objective function in exchange for the input X, the self-encoder is trained not to reconstruct its input; The output vector has the same dimensions as the input vector X; It means that the number of existing neurons in the input and output layers is equal. In this network, the output is the reconstruction of the input, and the error back-propagation algorithm is used for learning. Self-encoders teach network reconstruction by minimizing the error. The amount of existing neurons in the hidden layer is less than the encoding and decoding layers. In this study, a deep self-encrypting neural network is used to classify processed data. Deep learning models can acquire complex and non-linear dependencies from raw data and optimize the features to maximize performance. One of the advantages of using this type of networks is that they will not need special pre-processing compared to common models such as artificial neural networks, and it will automatically train the characteristics of raw and noisy data. Deep learning is more accurate and powerful compared to machine learning, if its components cannot be fully explained. The self-encrypting neural network is one of the unsupervised

networks that does not use probability distribution for training and instead has a goal like supervised networks and is trained with gradient descent and error back propagation. And it performs well in data classification and becomes more accurate. article, taking into account the presence of self-encrypting deep neural networks, the network tries to simulate the inputs to the output by encoding the inputs. The alternative form of learning in this network will be done in a supervised manner. Forcing the network to find models similar to the input data reduces the error in this work. The generated codes of the hidden layers of the network are the extracted features. These obtained features are classified and identify existing classes. Using this method increases processing speed and reduces costs related to memory.

3-1-2- The second scenario of the proposed method

The only difference between this scenario and the first proposed scenario is in the step of reducing its dimensions. In this scenario, instead of using bat meta-heuristic algorithm, principal component analysis method is used. The purpose of this work is to investigate the effect and power of statistical methods as well as meta-heuristics on the detection power of neural network and its execution time.

The steps of this algorithm to reduce dimensions are as follows:

First step: data collection

In this step, the data set obtained from the simulation is arranged in the form of a matrix and is prepared in order to reduce the feature.

The second step: adjusting the data

Adjusting the data means obtaining the average coordinates of the total data in each dimension and subtracting this obtained value from the coordinates of the total data in that dimension.

The third step: evaluation of the covariance matrix

In this step, the covariance matrix between the data coordinates in each dimension with the data coordinates in the same dimension and in all other dimensions is obtained. If the data set contains n dimensions, the covariance matrix of an n*n matrix is as follows:

$$cov_{n\times n} = \begin{bmatrix} cov(dim_1, dim_1) & .. & cov(dim_1, dim_1) \\ cov(dim_n, dim_1) & .. & ov(dim_n, dim_n) \end{bmatrix} \quad (2)$$

In the above equation, $dim_1, \dots, dim_n$ mean the different dimensions of the signals obtained from the measurement.

Step four: Evaluation of values and special vectors of the covariance matrix

This step is done in order to obtain the eigenvalues for the covariance matrix and create the eigenvector matrix.

The eigenvalues for the purpose of the cov matrix are evaluated based on the following relationship:

$$\det(\lambda I - cov) = 0 \quad (3)$$

In the above example, λ is the eigenvector matrix and I is the same matrix. After that, the eigenvector corresponding to each of the eigenvalues is evaluated through the following relationship:

$$(\lambda_i I - cov)V_i = 0 \quad (4)$$

In the above equation, V_i means the eigenvector corresponding to the desired eigenvalues.

Step Five: Selecting the parameters and constructing the feature vector

What is finally evaluated is a matrix containing n eigenvectors. Since each of the eigenvectors corresponding to larger eigenvalues has a greater effect on the ability to separate information, and as usual, the first m eigenvector is selected in the eigenvector matrix and is obtained based on the equation of the subset of the final dataset obtained from data transformation :

$$FinalData = RowFeatureVector * RowAdjustData \quad (5)$$

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

In the above equation, Final Data means the matrix obtained from the initial data transformation, and Row Feature Vector means the matrix of special selection vectors, and finally, RowAdjustData means the matrix of the initial data set.

4- Results

In this part, the simulation results of the proposed method will be presented. For this, two scenarios were defined. In one scenario, we will reduce the features by using the principal component analysis method, and then we will classify and recognize using the self-encrypting neural network. In the second scenario, to compare the accuracy and execution time of the algorithm, we reduced the dimensions using the bat evolutionary algorithm and gave the selected features to the self-encrypting neural network. At the end of this section, the results of both feature reduction methods are compared in terms of accuracy and execution time.

4-1- Database

,, the database related to the identification of different types of attacks, which is NSLKDD, has been used. The information of this database is as follows.

The database in question contains a large amount of normal and attack traffic that was obtained in a simulation on the local network of the US Air Force over a period of 9 weeks.

4-2- Evaluation criteria and parameters

The evaluation criterion is very important in order to determine the best method for conducting the test. ,, the criteria of accuracy, correctness, coverage and F1 are used for this purpose, and its relationship is as follows.

- **Accuracy criterion**

$$Accuracy = \frac{\sum True\ Positive + \sum True\ Negative}{TP+TN+FP+FN} \quad (6)$$

In the above relationships, TP is correct positive, TN is correct negative, FN is false negative and FP is false positive.

The accuracy measure, as shown in the above section, displays the percentage of information that is correctly classified and is introduced under the title of overall accuracy.

- **Validity criterion**

$$Precision = \frac{\sum TP}{\sum Test\ Outcome\ Positive} \quad (7)$$

- **Coverage criteria**

$$Recall = \frac{\sum True\ Positive}{\sum Realy\ Positive} \quad (8)$$

- **F1 Measure**

$$F1_{measure} = 2 \times \frac{Precision \times Pecall}{Precision + Pecall} \quad (9)$$

4-3- Simulation results of the first scenario (decrease dimension with bat algorithm)

In this scenario, after the data was called from the NSLKDD dataset, in the pre-processing step, the missing data were first normalized, and then the missing values were filled with zero. In the next step, the data were normalized and all the values of the features in the dataset were between 1 - They were placed up to 1.

In the next step, using the bat algorithm, we selected features and reduced the dimensions of less important features. In this algorithm, the initial population consists of 40 features in the data set, and each time the algorithm is executed, a subset of 30 is selected as the answer. Each set is validated using the evaluation function, which in this work is the accuracy of the neural network, a subset of features will be optimal,

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

which increases the accuracy more than other subsets. After obtaining the index of the selected features by applying them to the feature matrix, the unselected features were removed and the dimension of the feature matrix was reduced from 40 to 30.
In the following table (1) the parameters defined for the bat algorithm used in this scenario are displayed.

Table (1): Bat algorithm parameters

The goal of the algorithm	Find the best features
Primary population	100
Amounts per member	40 to the total number of features
stop condition	Reaching 50 repetitions or not progressing for 9 consecutive repetitions
Evaluation function	$cost_{fun} = 1 - Accuracy$
Number of selected features	30

After the features are selected. The data were divided into two groups, test and training. 70% for neural network training and 30% for evaluation. The structure of the used deep encoder neural network, including the layers, the number of neurons, and the activation functions can also be seen in Table (2).

Table (2): The structure of the deep-encoder neural network

The number of neurons in the input layer	30
The number of neurons of the first autoencoder layer	10
The number of neurons of the first autoencoder layer	10
The number of neurons in the output layer and the softmax layer	2
First and second dicoder conversion function	purlin
Compression adjustment factor (β)	0.001
The coefficient used in the cost function (ϒ)	4

After training to see the neural network with the selected features, the test data was also used to evaluate the neural network in the next step. The results obtained for this data set are given below. Figure (2) shows the confusion matrix for the method of feature selection with bat and classification with auto-encrypting neural network. In this matrix, label one reports security and label two reports penetration. The vertical axis of this matrix shows the output of the auto-encrypting neural network and the horizontal axis shows the real labels. This matrix states that 52.6% of the data predicted by the neural network correspond to the true label and are correctly recognized in class one, and also 46.8% of the data predicted by the neural network correspond to the true label and are correctly recognized in class two. are detected, and only 0.6% of the data are incorrectly detected in class one and two, and in the end it states that the overall accuracy for a single run is 99.4%.

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

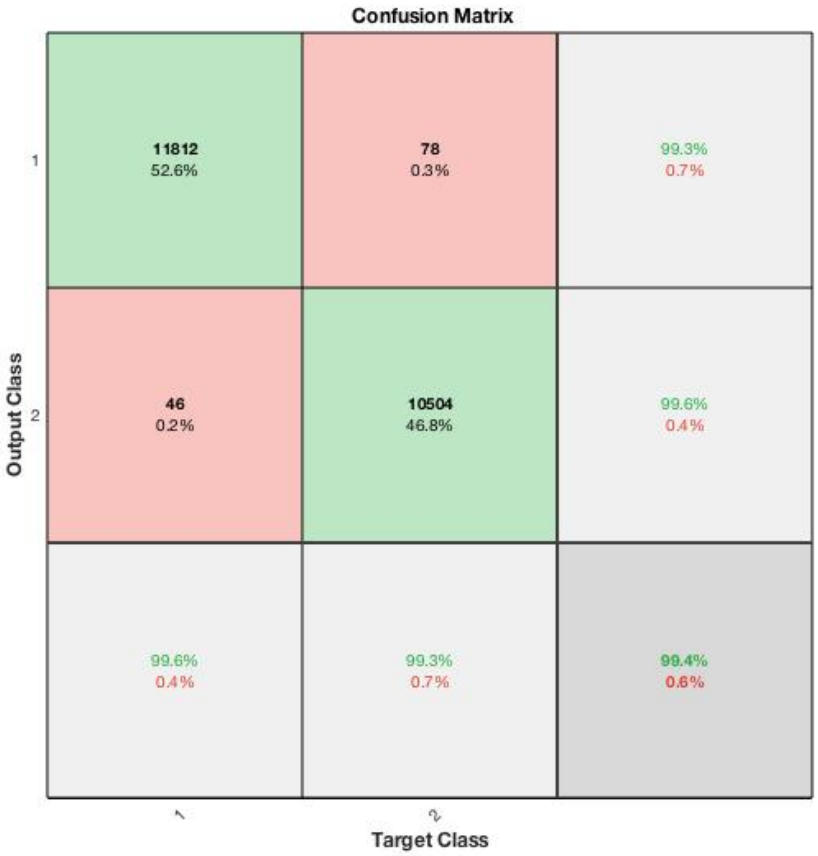

Figure (2): Confusion matrix for feature selection method with bat

Figure (3) shows the characteristic curve diagram for the first scenario, i.e. feature selection with the bat algorithm. This curve shows the rate of correct prediction versus wrong prediction of the autoencoder classifier. The higher this graph is above the bisector line, it shows that the ratio of the correct prediction of the classifier is higher than its incorrect prediction, so the detection accuracy is higher.

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

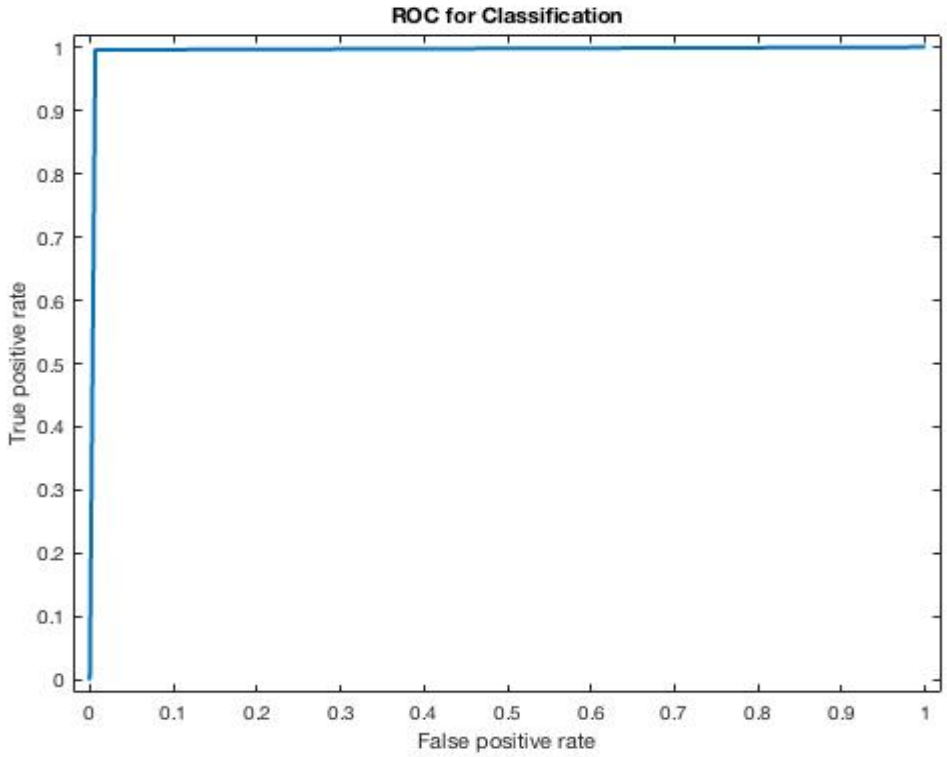

Figure (3): Performance characteristic curve for feature selection method with bat

In the final part of the results of the first scenario, the results obtained for the evaluation parameters are displayed as graphs in Figure (4).

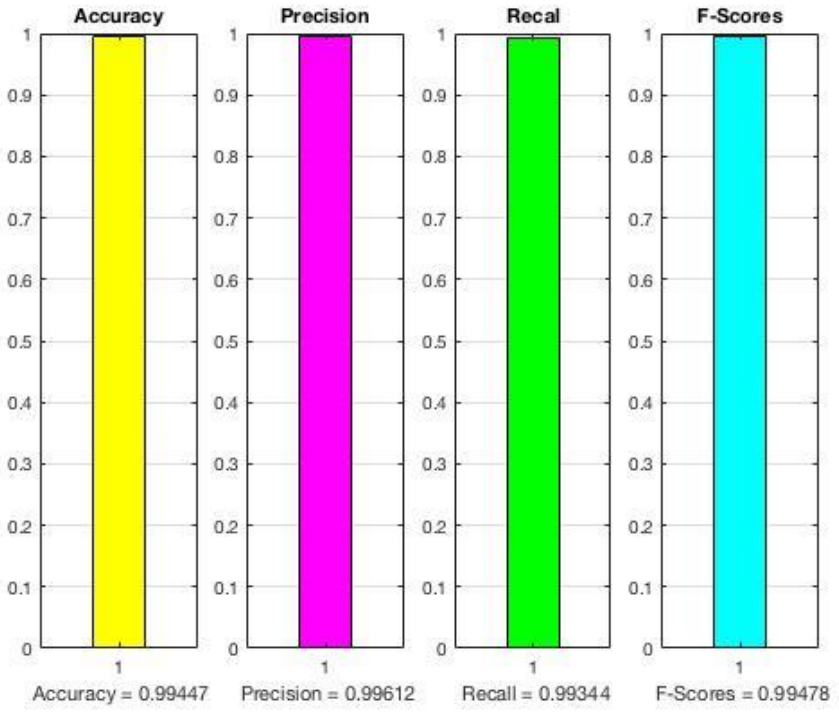

Figure (4): Results of evaluation criteria for the first scenario

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

The above figure shows that the accuracy obtained for this method is equal to 99.4%, the accuracy is equal to 99.6%, the coverage criterion is equal to 99.3%, and the F criterion is equal to 99.4% for a single execution of the pattern. It should be noted that the duration of this program is 17 It is minutes.

4-4- Simulation results of the second scenario (dimensional reduction with the principal component analysis algorithm)

As we stated in the introduction of this chapter, in this work, we intend to compare the results of dimensionality reduction using the principal component analysis algorithm with the meta-heuristic method in terms of accuracy and algorithm execution time. Therefore, in this scenario, all the conditions mentioned in the previous scenario for the pre-processing and the structure of the neural network of the autoencoder are the same. The only difference between this scenario and the previous method is to reduce the dimensions of the data from 40 dimensions to 30 dimensions using the principal component analysis method. The whole neural network structure for data classification is the same as the previous scenario. , we will express the results obtained by the dimension reduction method using the principal component analysis method.

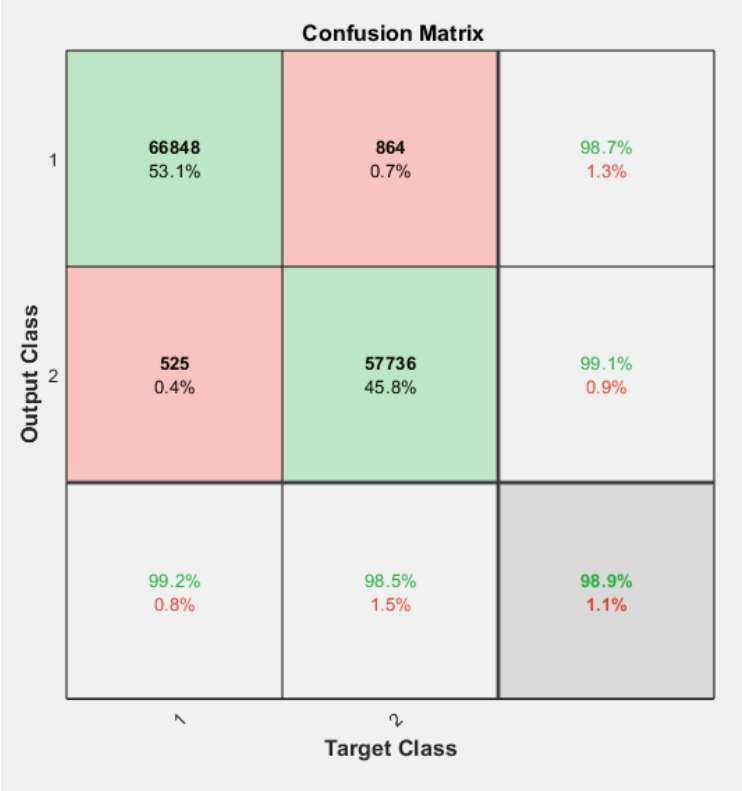

Figure (5): Confusion matrix for the selection method of the principal component baroush feature

Figure (5) shows the confusion matrix for feature selection method with principal component analysis method and classification with auto-encrypting neural network. This matrix shows that 53.1% of the data predicted by the neural network are in accordance with the true label and are correctly in class one. have been recognized and also 45.8% of the data predicted by the neural network are in accordance with the true label and have been correctly recognized in class two, and only 1.1% of the data have been wrongly recognized in class one and two and at the end it states The overall accuracy for one run is 98.9%.

Table (3) shows the results obtained in figures (6) and (7) for the auto-encoder principal component analysis method for the evaluation parameters.

Table (3): The resulting values for the principal component analysis method - self-encoder

True Positive	66848
True Negative	57736
False Negative	864
False Positive	525
Accuracy	0.9890
Recall	0.9872
Precision	0.9922
F_Score	0.9897

5- Conclusion

,, we discussed security in cloud computing. For this purpose and with the intention of increasing the accuracy of detection, two scenarios were defined. In the first scenario, after we pre-processed the data, we found suitable features using the bat algorithm, and by selecting useful features, we reduced the dimensions of the data and were able to achieve 99.4% accuracy. But the execution time of the algorithm in this scenario was nearly 17 minutes. In the other scenario, we performed dimension reduction using the principal component analysis algorithm and we were able to achieve 98.9% accuracy. The execution time in this scenario was 7 minutes. In both methods, data and classification had the same conditions. Although the accuracy of meta-heuristic methods was higher than the principal component analysis method, which is considered a statistical method, it is very important in runtime security detection systems. It can be concluded that in this application especially, due to the slight lower accuracy of the principal component analysis method compared to the bat meta-heuristic method, it is a better method, because it requires less time to detect intrusion or security.

References

[1] Krishnaveni, S., Sivamohan, S., Sridhar, S. S., & Prabakaran, S. (2021). Efficient feature selection and classification through ensemble method for network intrusion detection on cloud computing. Cluster Computing, 1-19.

[2] Prabhakaran, V., & Kulandasamy, A. (2021). Integration of recurrent convolutional neural network and optimal encryption scheme for intrusion detection with secure data storage in the cloud. *Computational Intelligence*, *37*(1), 344-370.

[3] Hizal, S., ÇAVUŞOĞLU, Ü., & AKGÜN, D. (2021, June). A New Deep Learning Based Intrusion Detection System for Cloud Security. In *2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)* (pp. 1-4). IEEE.

[4] Liu, Z., Xu, B., Cheng, B., Hu, X., & Darbandi, M. (2022). Intrusion detection systems in the cloud computing: A comprehensive and deep literature review. Concurrency and Computation: Practice and Experience, 34(4), e6646.

[5]Abdelsalam, M., Krishnan, R., Huang, Y., & Sandhu, R. (2018, July). Malware detection in cloud infrastructures using convolutional neural networks. In *2018 IEEE 11th International conference on cloud computing (CLOUD)* (pp. 162-169). IEEE.

[6] Anakath, A. S., Kannadasan, R., Joseph, N. P., Boominathan, P., & Sreekanth, G. R. (2022). Insider Attack Detection Using Deep Belief Neural Network in Cloud Computing. *COMPUTER SYSTEMS SCIENCE AND ENGINEERING*, *41*(2), 479-492.

[7] Kimmel, J. C., Mcdole, A. D., Abdelsalam, M., Gupta, M., & Sandhu, R. (2021). Recurrent Neural Networks Based Online Behavioural Malware Detection Techniques for Cloud Infrastructure. IEEE Access, 9, 68066-68080.
[8] Wang, W., Du, X., Wang, N.: Building a cloud IDS using an efficient feature selection method and SVM. IEEE Access 7 , 1345–1354 (2019)
[9] Ito, H., Kinoshita, Y., & Kiya, H. (2020, December). A framework for transformation network training in coordination with semi-trusted cloud provider for privacy-preserving deep neural networks. In 2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) (pp. 1420-1424). IEEE.