Iraqi Journal of Information and Communications Technology(IJICT)
Vol. 6, Issue 1, April 2023

www.ijict.edu.iq

ISSN:2222-758X
e-ISSN: 2789-7362

IMPROVING DETECTION FOR INTRUSION USING DEEP LONG SHORT-TERM MEMORY WITH HYBRID FEATURE SELECTION METHOD

Baraa I. Farhan[1], **Ammar D.Jasim**[2]
[1,2] College of Information Engineering, Al-Nahrain University, Baghdad, Iraq
bfarhan@uowasit.edu.iq[1],ammar @coie-nahrain.edu.iq [2]
Corresponding Author: Ammar D.Jasim
Received:29/5/2022; Revised: 31/07/2022; Accepted:01/10/2022
DOI:10.31987/ijict.6.1.213

***Abstract*- Due to the importance of the intrusion detection system, which is considered supportive of enhancing network security. Therefore, we seek to increase the efficiency of intrusion detection systems through the use of deep learning mechanisms. However, intrusion detection algorithms still suffer from problems in the process of classification and determining the presence and type of attack, which causes a decrease in the detection rate, an increase in the number of false alarms, and reduces system performance. This is due to a large number of redundant features that are not relevant to the dataset. To solve this problem, we propose a hybrid algorithm based on the use of the feature selection technique, which helps in reaching the goal optimally by choosing the best and most important features. It works by integrating three ways to reduce the number of features by deleting the static features that do not have much value from the information gained. This is done before the training stage by the deep learning model of LSTM as preprocessing for the CSE-CIC-IDS data set, which helps in improving the performance of the system by reducing the processing time and increasing the detection rate and accuracy. The results of the experiment showed a high accuracy of 99%**

keywords: **Intrusion Detection System (IDS), Deep learning (DL), Feature Selection, CSE-CIC-IDS2018, LSTM.**

I. INTRODUCTION

The wide growth in technology and applications and the emergence of the Internet of things have been accompanied by difficulties and security breaches [1]. Which requires the need to monitor the network and detect attacks and prevent intrusions, where an intrusion detection system plays an important role in overcoming security breaches [2]. By analyzing traffic and identifying suspicious activities and preventing them [3]. There are different methods for detecting intrusion in the network, the most common of which are signature-based detection, which relies on matching the signature of known attacks with traffic; and anomaly-based detection, which monitors the normal behavior of the network under normal conditions without attacks and distinguishes activities that deviate from normal behavior [4]. Deviation-based detection is better for its ability to monitor traffic and detect new attacks as a second level of protection in addition to firewall and authentication methods that prevent unauthorized access to the system [5]. The intrusion is detected in the network based on identifying the traffic by extracting its useful features and categorizing it into normal traffic or attack using a model that adopts one of the machine learning algorithms [6].

Given that the data of the Internet of Things is wide-ranging, it requires the use of deep learning, which is more efficient in dealing with large Datasets [7]. Especially the LSTM model for its ability to solve the vanishing scaling problem that appears in the model of traditional RNNs and linking communication records [8]. It is worth noting that many features affect the accuracy and detection time of the model, as it expands the scope of the search in IDS [9]. The use of feature

ISSN:2222-758X
e-ISSN: 2789-7362

selection methods is an excellent solution to exclude less relevant, fixed, repetitive, and useless techniques to reduce the dimensions of the data used to choose and classify the intrusion detection model, which helps improve performance [10], especially with data that contains a large number of features as in CSE-CIC-IDS2018. Feature selection models include three main types: the wrapper feature selection model, the filter model, and embedded technologies. Whereas the wrapper feature selection model evaluates the learning models to find ideal properties and takes a long computation time, which increases the possibility of getting the most lethal [11].

While in the embedded techniques model, the properties are selected in each iteration during the training phase and do not take a long time to calculate and help reduce the occurrence of the most lethal events [12]. In the filtering model, a special scale is used to make a subset of the features, and it is fast in calculating, and the percentage of getting the most lethal is low [13]. The filtering method based on the ratio of gain of information to intrinsic information is one of the best ways to improve deep learning models in addition to removing constant features and reducing dimensionality using constant features, quasi-constant features, and mutual information.

II. LITERATURE SURVEY

Given the importance of the topic of network traffic analysis and the detection of breaches in it, several recently published studies have addressed this topic. In addition to improving the business model through data reduction and feature selection, in this section, the most important works that mention this topic are listed.

Ismael R. et al. [3] proposed a system based on the use of a deep neural network (DNN) with the use of a feature selection method known as Binary Particle Swarm Optimization (BPSO) to improve the performance of the model. The performance of the model was tested on the CSE-CIC-IDS2018 dataset. The result of the model test showed an accuracy of 95%, faster processing, a good detection rate, and fewer false alarms.

Farhan, R. I. et al. [14] proposed a DNN-based intrusion detection system combined with a hybrid feature selection algorithm including two-particle optimized BPSO (BPSO) and correlation-based (CFS) to improve model performance. And it helped to solve the problem of selecting features efficiently and with 95% accuracy with little processing time and a high detection rate.

Alahmed S. I. et al. [15] proposed an intrusion detection system based on the use of generative adversarial networks (GANs) that helps provide better protection against adversarial perturbations. The Random Forest classifier was used, and feature selection methods such as principal component analysis (PCA) and recursive feature elimination (Rfe) were used to reduce data dimensions and enhance system resilience. The model was tested using the CSE-CICIDS2018 dataset, with an accuracy of 99.9%.

Laghrissi F. et al. [16] The paper presents the use of the LSTM model as a deep learning model that detects attacks with the use of feature selection and dimensionality reduction techniques represented by PCA (principal component analysis) and Mutual Information (MI) to improve the performance of the model. It was concluded that after applying the model to KDD99 data, PCA achieved a high accuracy rate. In the training and testing phases of binary and multiplayer ratings. Megantara,

ISSN:2222-758X
e-ISSN: 2789-7362

A. et al. [17] proposed a hybrid model for machine learning that combines a supervised model for feature selection and an unsupervised model for data reduction. It identifies the important features that are strongly relevant to the data using a decision tree, which helps to remove redundant features and distinguishes the local outlier factor technique. The model was tested on the NSL-KDD dataset and achieved a high accuracy of 99.89% in detecting R2L attacks. Ashiku, L. et al. [18] suggest an adaptive and flexible system for intrusion detection and classification. He focused on exploiting deep neural networks (DNNs) in facilitating IDS and discovering previously known and undiscovered features. Where, it has been tried to close the intrusion ports for intruders on the system and reduce penetration and test the system on a UNSW-NB15 dataset that represents real-time data to prove the accuracy of the model.

Lin et al. [19] viewed a proposed system to detect anomalies using LSTM long-term memory and Attention Mechanism (AM) to increase network training performance. The CIC-IDS 2018 data set has been used to train the proposed form, and the analysis of the results has mentioned the accuracy as 96.22%, the detection rate at 15% and the recall rate at 96%. Yuyang

Zhou et al. [20] proposed a framework that combines learning and feature selection that works for intrusion detection. A proposed CFS-BA algorithm was used to reduce the dimensions. Tested on three CIC-IDS2017 datasets, NSL-KDD and Aegean WiFi Intrusion Dataset (AWID), the accuracy rate of 99.9% and 99.5% has been mentioned.

III. METHODOLOGY

This section introduces the design of a hybrid model for network intrusion detection that uses deep learning techniques and feature selection methods as preprocessing stage, including constant features, quasi-constant features, and mutual information, which overcomes the high dimensions of network traffic content and reduces the features in the CSE-CIC-IDS2018 dataset that is used to test the model. This helps reduce training time and preserve the accuracy of the single LSTM model without pretreatment, which is 99.83

A. Long short-term memory (LSTM)

Deep learning is a more effective and accurate model for detection compared to machine learning models, especially with large and complex datasets [21]. It uses multiple hidden layers for processing that help increase accuracy and reduce costs by extracting features automatically instead of the method of feature engineering in machine learning (ML) [22]. The LSTM model represents the most important deep learning model as it solves the issue of long-term reliance problem that appears in the RNN model with serial data by distinguishing the current traffic and past traffic of the network and remembering it for a long time [23]. Whereas the hidden layer in RNN is simple and it is only tanh layer, while LSTM has four hidden layers as in Fig. 1, and it includes an input gate layer, an output gate layer, and a forget gate layer in addition to the main layer, and it has feedback connections [24],[25]. Table I shows the hyperparameters of the proposed LSTM model that help to avoid overfitting.

ISSN:2222-758X
e-ISSN: 2789-7362

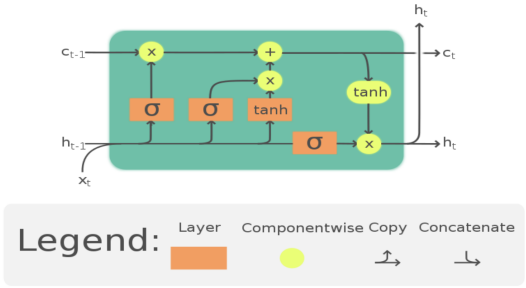

Figure 1: (LSTM) Architecture [25].

TABLE I
Hyperparameter of proposed LSTM model.

Parameters Name	Value
Hidden nodes in LSTM	150
Batch size	200
Epoch	30
The length of flow	10
Learning rate	0.001
Loss function	categorical crossentropy
Activation function	Relu, Soft max
Optimizer	Adam

B. Feature selection

The increase in data dimensions and the presence of repetitive and irrelevant data in the data set affect the performance of the learning model and reduce the efficiency of detection, classification, and prediction. Therefore, the use of dimension reduction mechanisms and feature selection represents a solution to this problem. It is implemented as a step before the training process, which helps reduce the complexity and time required for the training process. It is proposed in this research to use a hybrid method to reduce the dimensions and choose the features that are relevant and have the most relationships with the dataset, and it includes a constant feature, a quasi-constant feature, and utual information.

C. Constant Feature

The method of choosing features using constant features is one of the easiest types of filter methods used to delete fixed features. Where the values of these features do not show any difference for all recordings in the data set. These features are isolated and deleted as they are not useful in the training process, and take a lot of time. It is an easy and efficient way to reduce the dimensions of the data group and improve performance. 10 features were excluded and deleted by applying this method in our model to the CSE-CIC-IDS2018 data set because it is not useful in the training process.

ISSN:2222-758X
e-ISSN: 2789-7362

D. Quasi Constant Feature

One of the easiest filtering methods used to reduce dimensions and remove the Quasi-Constant feature. It is the method of selecting and removing Quasi Constant features as they are not useful for ranking and depend on the value of the threshold limit. Where if the minimum value specified for the threshold limit is 0.01 in our model applied to the CSE-CIC-IDS2018 dataset, it was found that only one feature was omitted because that feature was 99% of its recordings the same and had the same value from the dataset. The programmer can change the value of the threshold limit because it is used to determine the similarity ratio of the call to the semi-fixed feature and delete it. When the limit value was changed to 0.98 in our model, eight features were identified and omitted, as they were considered nearly the same. Changing the number of features omitted is due to a change in the threshold limit, which determines the measure of similarity. The Quasi-Constant feature method is an easy and effective way to reduce the dimensions of a data set and improve performance.

E. Mutual Information (MI)

The mutual information method is a method of calculating the statistical dependence between two variables and measures the amount of information for each feature in the data set. It helps to know the important features that affect the outcome, and they are the features that have a high degree of MI (which represents the amount of knowledge of one variable of uncertainty in another variable, and by increasing the value of MI, the uncertainty decreases). While it excludes features that have a low MI score or a value of zero, because that means that there is no relationship between these variables, which reduces the value of this feature. It is similar to the concept of correlation, but it is more general in that it does not represent a linear correlation. This method works better with discrete classes and values. The concept of MI is related to the concept of entropy of the random variable, which contributes to knowing the amount of information expected in the random variable. It is one of the basic ideas in information theory and is represented between the two variables Z | W | and is denoted I(Z; W). Cover and Tomas defined it[26]:

$$I(z;w)p_z(z) = \sum_{z,w} p_{zw}(z,w) \log \frac{p_{zw}(z,w)}{p_z(z)p_w(w)} = E_{P_{zW}} \log \frac{P_{zW}}{P_z P_w}$$
$$p_z(z) \text{ and } p_w(w) \text{ are the marginals: } p_z(z) = \sum_w P_{zw}(z,w) \tag{1}$$

ISSN:2222-758X
e-ISSN: 2789-7362

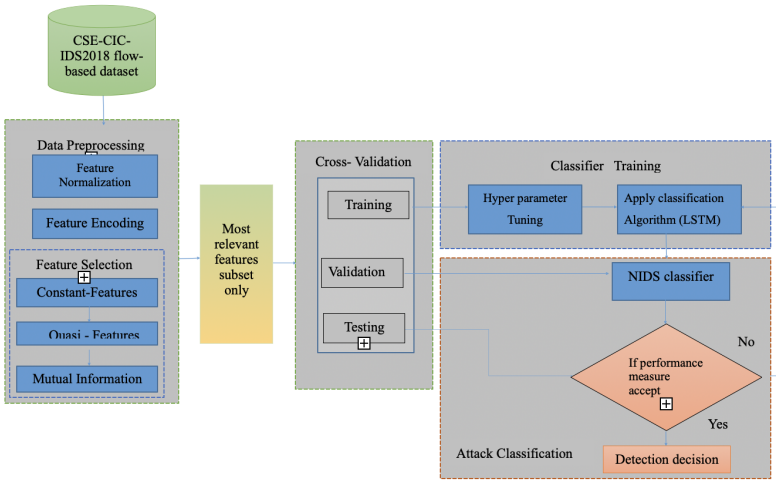

Figure 2: Flowchart for the proposed Model

F. Proposed Approach

In this section, we describe the proposed deep learning model represented by LSTM and the CSE CIC-IDS2018 dataset taken from the AWS platform. Which represents real data used in training and testing the model, where the data is divided into a training group and a test group as a step that precedes training and testing. The preprocessing steps that contribute to improving performance by encoding and normalizing features and removing static and useless features are performed using feature selection methods that include (Constant Feature, Quasi Constant Feature, and Mutual Information) methods, which help reduce the dimensions and features in the dataset, which helps increase efficiency and reduce time. Fig. 2 shows the flowchart for the proposed model.

G. Real Dataset (CSE-CIC-IDS2018)

The CSE-CIC-IDS2018 dataset is one of the most important real data sets used in the field of intrusion detection and represents a transformation from a static data set such as NSL-KDD to a dynamic dataset. This data is taken from the Amazon Platform (AWS) by Communications Security Corporation (CSE) and the Canadian Cyber Security Institute (CIC) and represents real-time network traffic [27]. It is considered one of the most reliable data sources for evaluating intrusion detection models based on network anomalies [14]. This data contains 16,000,000 instances collected in ten days and includes the latest attacks ten classes of attacks according to the percentage of detection in the data as shown in table II: Benign, Bot, FTP BruteForce, SSH-Bruteforce, DDOS attack-HOIC, DDOS attack-LOIC-UDP, DoS attacks - GoldenEye, DoS Attacks-Slow HTTP Test, Intrusion and Web attacks [28]. The original dataset contained 80 features. There are some features that have little effect on interpreting the behavior of data and traffic, whether it is normal or not. Therefore, these features, such as the timestamp feature and IP addresses, that do not help in training the neuron to detect errors and intrusions are deleted, so we use 78 features from the original number of features. It is divided into two types; one is

ISSN:2222-758X
e-ISSN: 2789-7362

normal, and the other is nine types of attacks classified according to the above-mentioned types as shown in table II with the percentages of each attack from the origin of the data.

TABLE II
Volume of data points in Attack Class and Ratio of it.

Class number	Attack Class	Volume of data points in class	Ratio from the original data(1252835 row)
1	Benign	971016	77.505%
2	Infilteration	38703	3.089%
3	DoS attacks-Hulk	37323	2.979%
4	Bot	137185	10.95%
5	DDOS attack-HOIC	57507	4.59%
6	DDOS attack-LOIC-UDP	8377	0.669%
7	FTP-BruteForce	2234	0.178%
8	DoS attacks-GoldenEye	332	0.026%
9	DoS attacks-SlowHTTPTest	103	0.008%
0	SSH-Bruteforce	55	0.004%

IV. Experimental Results and Discussion

The implemented hybrid model contains preprocessing by a hybrid feature selection method including (Constant Feature, Quasi Constant Feature, and Mutual Information) and uses the LSTM training model implemented by Visual Code 2019 on Python 3.9. It uses TensorFlow, which includes libraries (Panda, Scikit-Learn, and Numpy) and Keras in the Windows 10 environment. The model is applied to the preprocessed CIC-IDS 2018 data set, which is divided into two groups: 65% for training and 35% for testing, and then evaluating the results and performance of the model using evaluation scales.

A. Used Metrics to Evaluation

This paper used multiple metrics to evaluate the performance of the model; namely, the accuracy, loss, precision, recall, confusion matrix, and F1 Score are representations of the ability to classify samples correctly in a model.

$$\begin{aligned} \text{Accuracy} &= \frac{TP+TN}{TP+FP+FN+TN} \\ \text{Precision} &= \frac{TP}{TP+FP} \\ \text{Recall} &= \frac{TP}{TP+FN} \\ F1\ \text{Score} &= \frac{2+TP}{2*TP+NP+FN} \end{aligned} \tag{2}$$

1) True Positive (TP) is the correct classification of the attack as an attack.
2) True negative (TN) expresses that a normal input is properly classified as a normal input.
3) False Positive (FP) is incorrectly classifying a normal entry as an attack.
4) False negative (FN) expresses an attack incorrectly classified as a normal entry.

ISSN:2222-758X
e-ISSN: 2789-7362

Accuracy is the ratio of correctly classified samples to the total number of samples. Accuracy is inversely proportional to the false alarm rate (FAR). The higher the accuracy, the lower the false alarm rate. Fig. 3 shows the accuracy measurement in the training and testing phases. The loss function is the variation between the expected and actual output. Fig. 4 shows the loss measurement in the training and testing phases. The precision is the ratio of the predicted positive samples to the total number of positive samples. The recall represents the ratio of the predicted positive samples to the total number of samples. The confusion matrix is a graphical representation that summarizes the performance and accuracy of the classification process, illustrating true and false positive values and giving an idea of the errors that the model makes and how to correct them. where each row in the array represents the status of an expected class and each column represents the status of an actual class. F1 Score It is an important measure in the case of data with varying class average recall and precision. Predict natural and attacking packets in network traffic. Fig. 5 shows the confusion matrix.

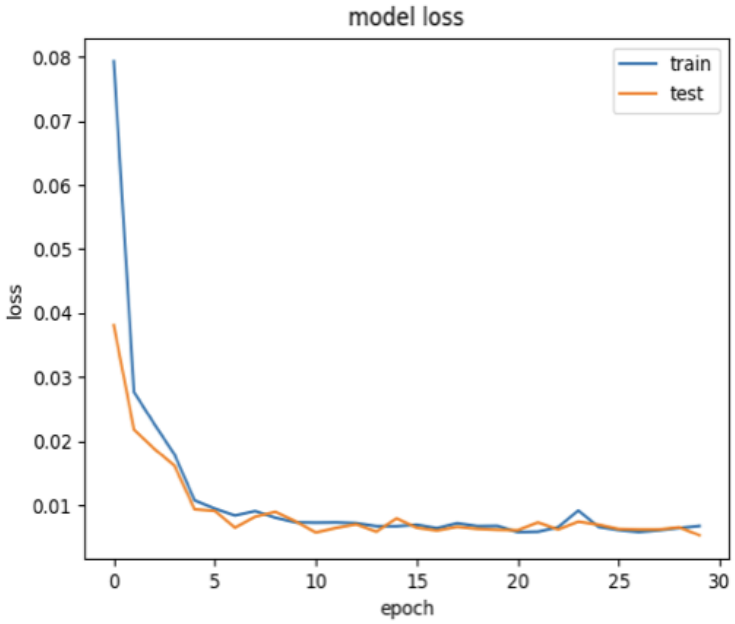

Figure 3: The Accuracy of Train stage and Test stage

ISSN:2222-758X
e-ISSN: 2789-7362

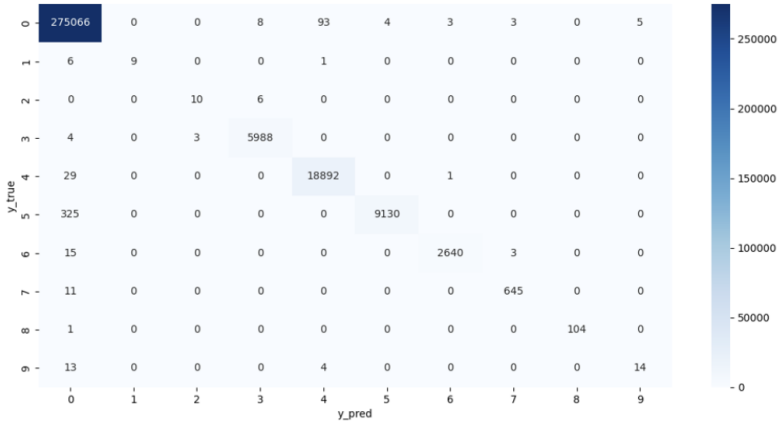

Figure 5: Confusion Matrix

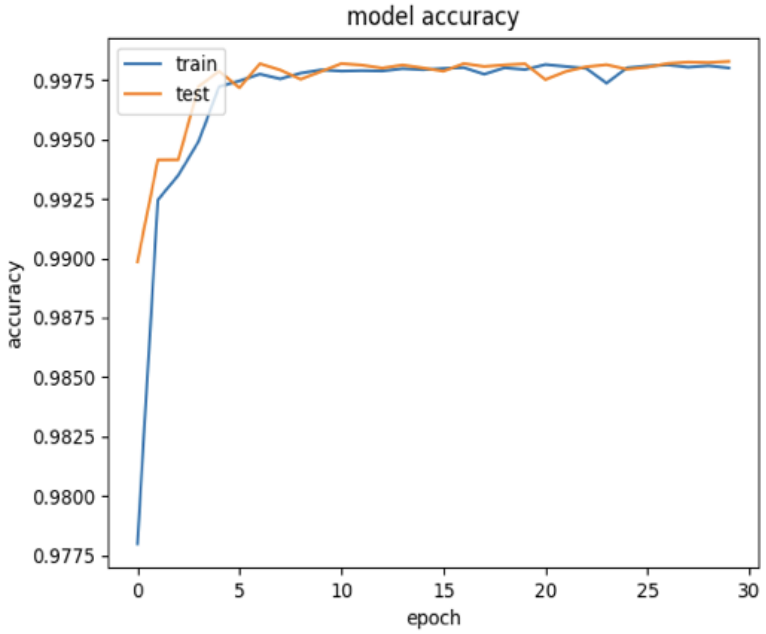

Figure 4: Loss of Train stage and Test stage

V. Comparative Analysis

In this part, the difference between our LSTM model with hybrid features selection method and other previous models is shown in Table III. This is used as a preprocessing to reduce the features and choose the most important features, where 50 features were selected out of 80 features in the data set, which reduces training time.

TABLE III
The Comparison Among our Model and Another Method

Research	Feature selection	DL, ML	Data set	.
Ismael R. et al. (2020) [3]	BPSO	DNN	CSE -CIC-IDS 2018	90.25%
Farhan, R. I. et al. (2021) [14]	CFS, BPSO	DNN	CSE -CIC-IDS 2018	95%
Alahmed S, et al. (2022) [15]	PCA, Rfe	GANs	CSE -CIC-IDS	99.9%
The proposed model	Const- Feature, Quasi Const- Feature, MI	LSTM	CSE -CIC-IDS	99.83%

Our used model showed an accuracy of 99.83%, which is better than the rest of the previous models. The use of a hybrid model consisting of three methods for selecting features helped improve the model. The learning model is represented by LSTM, which is considered one of the best learning models.

VI. Conclusion

In this paper, a network intrusion detection system using deep learning technology is proposed. Where the LSTM model was applied to build the neural network. To improve and support the performance of the model, a pretreatment consisting of a hybrid method was used that combines three types of feature selection methods (Const-Feature, Quasi Const-Feature, and MI). It defines the most relevant and non-redundant features that support the detection method applied to the CSE-CIC-IDS2018 real dataset. This helped to maintain good accuracy of up to 99.83%, reduce errors, and speed up the training process by reducing the number of features from 80 to 50. Finally, looking to the future, we plan to use a multi-layer model to increase detection and to test the model on another dataset to support the validity and suitability of the model.

Funding

None

ACKNOLEDGEMENT

The author would like to thank the reviewers for their valuable contribution in the publication of this paper.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

References

[1] D. Chen et al., "Privacy-Preserving Encrypted Traffic Inspection with Symmetric Cryptographic Techniques in IoT," in IEEE Internet of Things Journal, doi: 10.1109/JIOT.2022.3155355.
[2] Baraa I. Farhan, Ammar D. Jasim. Performance analysis of intrusion detection for deep learning model based on CSE-CIC-IDS2018 dataset, IJEECS, Vol 26, No 2, Pages: 1165-1172, doi: http://doi.org/10.11591/ ijeecs.v26.i2.pp %25p, 2022.
[3] Ismael R, Abeer F, Nidaa TM, F. Optimized Deep Learning with Binary PSO for Intrusion Detection on CSECIC-IDS2018 Dataset. J Al Qadisiyah Comput Sci Math. 2020;p. 16-16, DOI: https://doi.org/10.29304/ jqcm.2020.12.3.706
[4] G. Pang, C. Shen, L. Cao, and A. V. D. Hengel. Deep learning for anomaly detection. ACM Computing Surveys (CSUR), vol. 54, no.2, pp. 1-38, 2021. https://doi.org/10.1145/3439950.
[5] Baraa I. Farhan, A. D.Jasim. A Survey of Intrusion Detection Using Deep Learning in Internet of Things. Iraqi Journal For Computer Science and Mathematics, vol. 3, no. 1, pp. 81-91, Jan. DOI: https://doi.org/ 10.52866/ ijcsm.2021.02.02.002, 2022.

[6] V. Kanimozhi, T. Prem Jacob,Artificial Intelligence outflanks all other machine learning classifiers in Network Intrusion Detection System on the realistic cyber dataset CSE-CIC-IDS2018 using cloud computing, ICT Express, Volume 7, Issue 3,2021,Pages 366-370,https://doi.org/10.1016/j.icte.2020.12.004 https://doi.org/10.1016/ j.icte.2020.12.004.
[7] M. Abdel-Basset, H. Hawash, R. K. Chakrabortty and M. J. Ryan, "Semi-Supervised Spatiotemporal Deep Learning for Intrusions Detection in IoT Networks," in IEEE Internet of Things Journal, vol. 8, no. 15, pp. 12251-12265, 1 Aug.1, 2021, doi: 10.1109/JIOT.2021.3060878.
[8] X. Zhou, Y. Hu, W. Liang, J. Ma, and Q. Jin, "Variational LSTM enhanced anomaly detection for industrial big data", IEEE Transactions on Industrial Informatics, vol. 17, no. 5, pp. 3469-3477, 2020. https://doi.org/10.1109/ TII.2020.3022432.
[9] X. Zhou, Y. Hu, W. Liang, J. Ma, and Q. Jin, "Variational LSTM enhanced anomaly detection for industrial big data", IEEE Transactions on Industrial Informatics, vol. 17, no. 5, pp. 3469-3477, 2020. https://doi.org/10.1109/ TII.2020.3022432.
[10] Q. R. S. Fitni and K. Ramli, "Implementation of Ensemble Learning and Feature Selection for Performance Improvements in Anomaly-Based Intrusion Detection Systems," 2020 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT), 2020, pp. 118-124, doi: 10.1109/ IAICT50021.2020.9172014.
[11] Antonia Nisioti, Alexios Mylonas, Paul D. Yoo, Vasilios Katos, "From Intrusion Detection to Attacker Attribution: A Comprehensive Survey of Unsupervised Methods", DOI: 10.1109/COMST.2018.2854724, IEEE.
[12] Murooi Khalid Ibraheem, et al., " Network Intrusion Detection Using Deep Learning Based On Dimensionality Reduction", REVISTA AUS 26-2, DOI:10.4206/ Aus. 2019.n26.2.23.
[13] Liu Z, Thapa N, Shaver A, Roy K, Siddula M, Yuan X, Yu A. Using Embedded Feature Selection and CNN for Classification on CCD-INID-V1-A New IoT Dataset. Sensors. 2021; 21(14):4834. https://doi.org/10.3390/ s21144834.
[14] Farhan, R. I. ., Maolood, A. T. ., & Hassan, N. . (2021). Hybrid Feature Selection Approach to Improve the Deep Neural Network on New Flow-Based Dataset for NIDS. Wasit Journal of Computer and Mathematics Science, 66-83. https://doi.org/10.31185/wjcm.Vol1.Iss1.10.
[15] Alahmed S, Alasad Q, Hammood MM, Yuan J-S, Alawad M. Mitigation of Black-Box Attacks on Intrusion Detection Systems-Based ML. Computers. 2022; 11(7):115. https://doi.org/10.3390
[16] Laghrissi, F., Douzi, S., Douzi, K., Hssina, B. Intrusion detection systems using long short-term memory (LSTM). Journal of Big Data, 2021, 8(1), 1-16, https://doi.org/10.1186/s40537-021-00448-4.
[17] Megantara, A. A., & Ahmad, T. A hybrid machine learning method for increasing the performance of network intrusion detection systems. Journal of Big Data, 2021, 8(1), 1-19. https://doi.org/10.1186/s40537-021-00531-w.
[18] Ashiku, L., Dagli, C. Network Intrusion Detection System using Deep Learning. Procedia Computer Science, $2021, 185, 239 - 247$. https://doi.org/10.1016/j.procs.2021.05.025.
[19] $P.Lin, K.Ye, and C. - Z.Xu, "Dynamic network anomaly detection system by using deep learning techniques, "in Cloud Computing - CLOUD2019, vol.11513 LNCS, 2019, pp.161 - 176. DOI : 10.1007/978 - 3 - 030 - 23502 - 4_12.$
[20] Yuyang Zhou, Guang Cheng, Shanqing Jiang, Mian Dai,Building an efficient intrusion detection system based on feature selection and ensemble classifier,Computer Networks, Volume174,2020,107247,https://doi.org/10.1016/ j.comnet.2020.107247.
[21] H. S. Abdulkarem and A. D. Alethawy, "DDOS attack detection and mitigation at sdn enviroment," Iraqi Journal of Information and Communications Technology, vol. 4, no. 1, pp. 1-9, 2020. DOI: 10.1109/ GPECOM49333.2020.9247850.
[22] Yuyang Zhou, Guang Cheng, Shanqing Jiang, Mian Dai,Building an efficient intrusion detection system based on feature selection and ensemble classifier,Computer Networks, Volume174,2020,107247,https://doi.org/10.1016/ j.comnet.2020.107247.
[23] Yuyang Zhou, Guang Cheng, Shanqing Jiang, Mian Dai,Building an efficient intrusion detection system based on feature selection and ensemble classifier,Computer Networks, Volume174,2020,107247,https://doi.org/10.1016/ j.comnet.2020.107247.
[24] HakanAydin, Zeyneporman, Muhammed AliAydin, "A Long Short-Term Memory (LSTM)-Based Distributed Denial of Service (DDoS) Detection and Defense System Design in Public Cloud Network Environment", Computers & Security, 2022, https://doi.org/10.1016/j.cose.2022.102725.
[25] Laghrissi, F., Douzi, S., Douzi, K. et al. Intrusion detection systems using long short-term memory (LSTM). J Big Data 8, 65 (2021). https://doi.org/10.1186/s40537-021-00448-4.
[26] Cover TM, Thomas JA. Information theory and statistics. In: Elements of information theory. 2nd ednn. Wiley; 2005. https://doi.org/10.1002/047174882X.ch11.
[27] Zhou, Y., Cheng, G., Jiang, S., & Dai, M., "Building an efficient intrusion detection system based on feature selection and ensemble classifier". Computer Networks, doi:10.1016/i.comnet.2020.107247.
[28] Registry of Open Data on AWS, datasets that are available via AWS resources, Accessed on: May 30,2021. [Online]. https://registry.opendata.aws/cse-cic-ids2021/.

