Intrusion Detection System for NSL-KDD Dataset Based on Deep Learning and Recursive Feature Elimination

Article *in* Engineering and Technology Journal · September 2021

DOI: 10.30684/etj.v39i7.1695

2 authors, including:

Ekhlas K. Gbashi
University of Technology-Iraq
36 PUBLICATIONS **152** CITATIONS

Engineering and Technology Journal
Journal homepage: engtechjournal.org

Intrusion Detection System for NSL-KDD Dataset Based on Deep Learning and Recursive Feature Elimination

Bilal Mohammed [a*], **Ekhlas K. Gbashi** [b]

[a] Computer science, University of Technology, Baghdad, Iraq, cs.19.65@grad.uotechnology.edu.iq

[b] Computer science, University of Technology, Baghdad, Iraq, 110026@uotechnology.edu.iq

*Corresponding author.

Submitted: 26/04/2020 Accepted: 10/08/2020 Published: 25/07/2021

K E Y W O R D S

Classification, Intrusion Detection Systems, NSL-KDD, Deep Neural Network, Recurrent Neural Network.

A B S T R A C T

Intrusion detection system is responsible for monitoring the systems and detect attacks, whether on (host or on a network) and identifying attacks that could come to the system and cause damage to them, that's mean an IDS prevents unauthorized access to systems by giving an alert to the administrator before causing any serious harm. As a reasonable supplement of the firewall, intrusion detection technology can assist systems to deal with offensive, the Intrusions Detection Systems (IDSs) suffers from high false positive which leads to highly bad accuracy rate. So this work is suggested to implement (IDS) by using a Recursive Feature Elimination to select features and use Deep Neural Network (DNN) and Recurrent Neural Network (RNN) for classification, the suggested model gives good results with high accuracy rate reaching 94%, DNN was used in the binary classification to classify either attack or Normal, while RNN was used in the classifications for the five classes (Normal, Dos, Probe, R2L, U2R). The system was implemented by using (NSL-KDD) dataset, which was very efficient for offline analyses systems for IDS.

How to cite this article: Bilal Mohammed, Ekhlas K. Gbashi , "Intrusion Detection System for NSL-KDD dataset based on deep learning and recursive feature elimination," Engineering and Technology Journal, Vol. 39, No. 07, pp. 1069-1079, 2021.
DOI: https://doi.org/10.30684/etj.v39i7.1695

1. INTRODUCTION

Intrusion detection systems is a security technique which analyses network systems and computer in real time to detect intrusions and manage responsive actions [1]. Signature and Anomaly are two major models that utilized in intrusion detection systems. The anomaly is depending on the statistical description of programs or users which is mean detecting any activity deviating from the profile of normal behavior The Signature-based IDSs depends on gathering and saving the signature of known attacks in database. [2,3]. The very huge increase in networks and with the increase of devices and users, the computers suffer from attacks and security vulnerabilities which be difficult and expensive to be solved, so the best solution is intrusion detection system to control and monitor the traffic in network. The paper revision for anomaly detection fully based on deep machine learning ways on

different training and testing dataset. The suggested system was implemented by NSL-KDD dataset because of the problems in KDD99 dataset [4]. The Network Intrusion Detection System (NIDS) is estimate to know attacks which need a comprehensive data set that contains known and unknown behaviors [5]. This work proposes a network intrusion detection system which depends on deep neural network and recurrent neural network to classify the normal and the attacks. The paper is arranged as follows: section two illustrates some related works, section three shows Descriptions for the dataset as inclusive and substantive, section four and five illustrates Recurrent and deep neural networks, section six explains the evaluation metrics, section seven explain the steps of suggested system, section eight illustrates the experimental results and discussions and finally conclusions and future work.

2. RELATED WORK

Many recent systems depend on normal machine learning techniques. One public ways to build intrusion detection systems is to utilize Artificial Neural Network . Such as the back-propagation algorithm [6].

Many of the common methods are used to detect intrusion in intrusion detection systems, such as Support vector machines [7,8], K-nearest neighbor (KNN)[9], and Random forest (RF) [10]. [11] Suggested IDS on the NSL-KDD dataset by using the support vector machine and decision tree algorithms [12], respectively. Use anomaly detection with the Naive Bayes (NB)[13] and examined on the KDD99 [14]. Examined DARPA dataset by utilized Support Vector. Examine of KDD dataset by added the C4 and Self-Organization Map [15].

Albeit, to make some perfection, such as the accuracy and decreasing the size of false alarms [16] . Based on the KDD dataset and using Long short term memory algorithm for feature selection and classify an attacks in dataset [17].

Machine learning learn the particulars of TCP/IP attributes, but Deep learning is a part of machine learning that be complex because it consists of many layers and transit the TCP/IP in many layers. design this model that combine discretization and HNB classifier this approach focused on problems in intrusion detection and this model based on hidden layer in NB model for many class than can get better accuracy and high detection rate of attacks [18].

Newly, intrusion detection systems became on deep learning ways . In [19],suggested for intrusion detection a proposal Self-taught Learning (STL) . According to [20], LSTM assemble many-features using various sources features, such as numeric, nominal, and binary features. In [21], the authors detected how to establish an (IDS) with RNN ..

3. DATASET DESCRIPTION

Since 1999, the Knowledge Discovery and Data Mining (KDD'99), wildly used data set for the estimated of anomaly detection techniques [22]. Some investigators examine KDD99 but results were poor execution on the anomaly detection approach so the best solution was using new dataset which is NSL-KDD [23]. This dataset consists of chosen records of the all KDD data set. The Training dataset consists of (125,973) and test dataset consists of (22,544) samples each of sample contains 41 attributes either attacks or normal. Attacks in this dataset are split into: Root to Local (R2L), User to Root(U2R), Denial of service(DOS), Probe Attacks. The following shows some explanations for these attacks types: -

1. Dos: It is a class of attack where the attacker restricts processing time of the resources so as to avoid the real user from obtaining those resources.
2. R2L: Attackers are not allowing access from a remote system.R2L attack is both categorized under network based and host-based (NIDS).
3. U2R: the attacker tries to gain the password of the user and then get into the system as a legitimate user and retrieve the data.
4. Probe: The Attacker will exam the network to collect information and would make some penetrations in the next time.
5.

In addition to these types we also have a class describing the Normal class. Table I shows the four

types of attack in NSL_ KDD [23].

TABLE I: Attacks in NSLKDD Dataset

Intrusions	Intrusions Type
DOS	Back,Smurf,Teardrop,Pod,ProcesstableLand,Apache,Mailbomb,Udpstorm,Neptune, Worm.
probe	Mscan,IPsweep,Satan,Saint ,Portsweep, Nmap.
R2L	Sendmail,Xsnoop,Ftp_write,hop,Imap,MultiNamed,Snmpgetattack,Guess_password,Phf,Snmpgue,Xlock,Httptunnel,Warezmaster,ss.
U2R	Xterm,Rootkit,Loadmodule,Ps,Buffer_overflow,Perl,Sqlattack

4. RECURRENT NEURAL NETWORKS

Recurrent neural networks are a type of supervised Deep learning models, made of artificial neurons with one or more returns loops. The returns loops are recurrent rotation over time or sequence [24]. A Recurrent neural network has been successfully used for text data, speech data, classification, regression, natural language processing and generative models [25]. A recurrent neural networks not suitable for image data and tabular data.

5. DEEP NEURAL NETWORK

A deep neural network is a neural network with a certain level of complexity, this algorithm has more than two layers. Deep neural networks utilize sophisticated Mathematical modeling to process data in complex methods. A deep neural network (DNN) has been successfully utilized for a number of classification and regression systems including image classification and natural language processing. It also used for a speech recognition [26] and Bayesian speech and language processing [27].

6. EVALUATION METRICS

Because intrusion detection systems performance based on Confusion Matrix to evaluate classification in the actual and predicted as shown in table II.

- True Positive (TP): -the model correctly predicted normal as normal.
- True Negative (TN): - the model correctly predicted attacker as attacker.
- False Positive (FP): - the model incorrectly identify a normal activity as a malicious one.
- False Negative(FN): - the model incorrectly identify malicious traffic as normal

TABLE II: Shows the Confusion Matrix.

		Actual Class	
Predicted Class		**Normal**	**Attacker**
	Normal	TP	FP
	Attacker	FN	TN

The following metrics are the most usually utilized evaluation metrics: -

1) Accuracy:

Defined as the average of truly classified samples as normal or attack over the whole number of samples.

$$Accuracy = \frac{TP + TN}{Total\ Number\ of\ Samples} \times 100 \quad (1)$$

2) Precision (P):

Is the proportion of positive predictions made by the classifier that are true, as in the following equation.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

3) Recall*(R):*

It is the percentage of correct positive that is truly detected by the classifier and called DR, TPR.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

4) F1-Score(F1):

Is the result between of the precision and the recall**.**

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

5) Receiver Operating Characteristics (ROC) curve

ROC: - is graph based on the True Positive Rate (TPR) to False Positive Rate (FPR) . The machine learning model was better if the AUC is higher.

$$AUC = \int_0^1 \frac{TP}{TP+FN}\, d\, \frac{FP}{TN+FP} \quad (5)$$

7. THE PROPOSED SYSTEM

Based on NSL-KDD dataset, the suggested system for designing an Network Intrusion Detection System. After applying preprocessing on raw dataset. Normalization is applied to make the values of all features are between 0 and 1.Training set was used to train model and testing set was used to evaluate the trained model. To select important features from Training set we applying RFE technique and these important features will be used with testing set at prediction part. After select features for training set, a classification is implemented on training set by using Deep Neural Network and Recurrent Neural Network algorithms. Finally the model is evaluated by prediction with test set and compare results. The proposed system is shown in Figure 1.

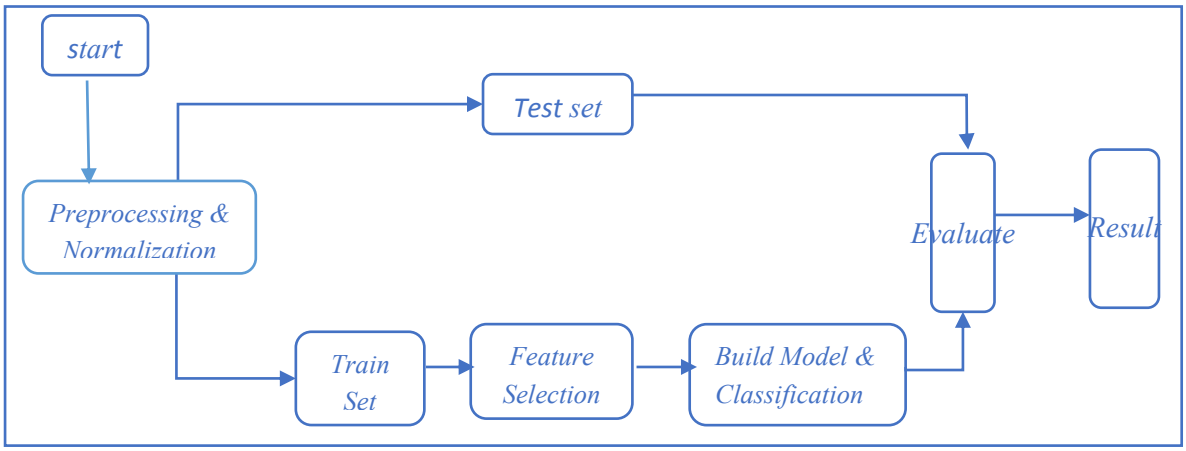


Figure 1: The proposed System

I. Pre-processing: -

Recurrent and deep Neural networks based classification uses only numerical values for training and testing. Hence a pre-processing stage is must to convert the non-numerical values to numerical values. Two main tasks in our pre-processing are:

- The nominal attributes in the dataset converting to numerical values. The features 2, 3 and 4 are the protocol type, service and flag.
- The attack types at the end of the dataset convert into its numeric categories.

II. Normalization: -

The features in the NSL-KDD dataset is continuous or discrete values. The ranges of the values were different and this made them incomparable. The features were normalized by subtracting mean from all feature and dividing by its standard deviation, then normalized the test features using the mean and standard deviation of each feature from train datasets. Min-Max normalization way which is a linear transformation is utilized to scale data between (0,1).The following method is utilized to find the new value [28]:

$$X_n = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (6)$$

III. Features Selection: -

One of the essential phases in data pre-processing for intrusion detection system is feature selection. It decreases the number of features, split redundant, irrelevant, or noisy data, and fetch the essential features that effects for intrusion detection system. The wrapper based Recursive Feature Elimination for Random Forest Classifier (RFC-RFE) method, is utilized as feature selection technique which represent wrapper selection way. Recursive Feature Elimination (RFE) provides feature weights that believe multivariate reacting effects between features. RFE was proposed in the state of Random Forest Classifier for getting the best subset from features. To find the best attribute subset, without of doing an exhaustive test over whole feature sets, RFE uses Wrapper technique by using the particular model (Random Forests) and reject the bad Feature (by absolute classifier weight or attribute ranking), and reiteration the operation over increasingly minimal attribute subsets while the best model hypothesis is completed. The weights of this good model are utilized to rank features. Utilized the sickie-learn achievement of RFE with random forest to come up with a feature ranking for dataset. In the NSL-KDD, 25 of important features were selected. As in the form that contains successive operations on the algorithm for feature selection, has achieved a high score reach more than of 98%. Figure II shows the important features resultant from REF.

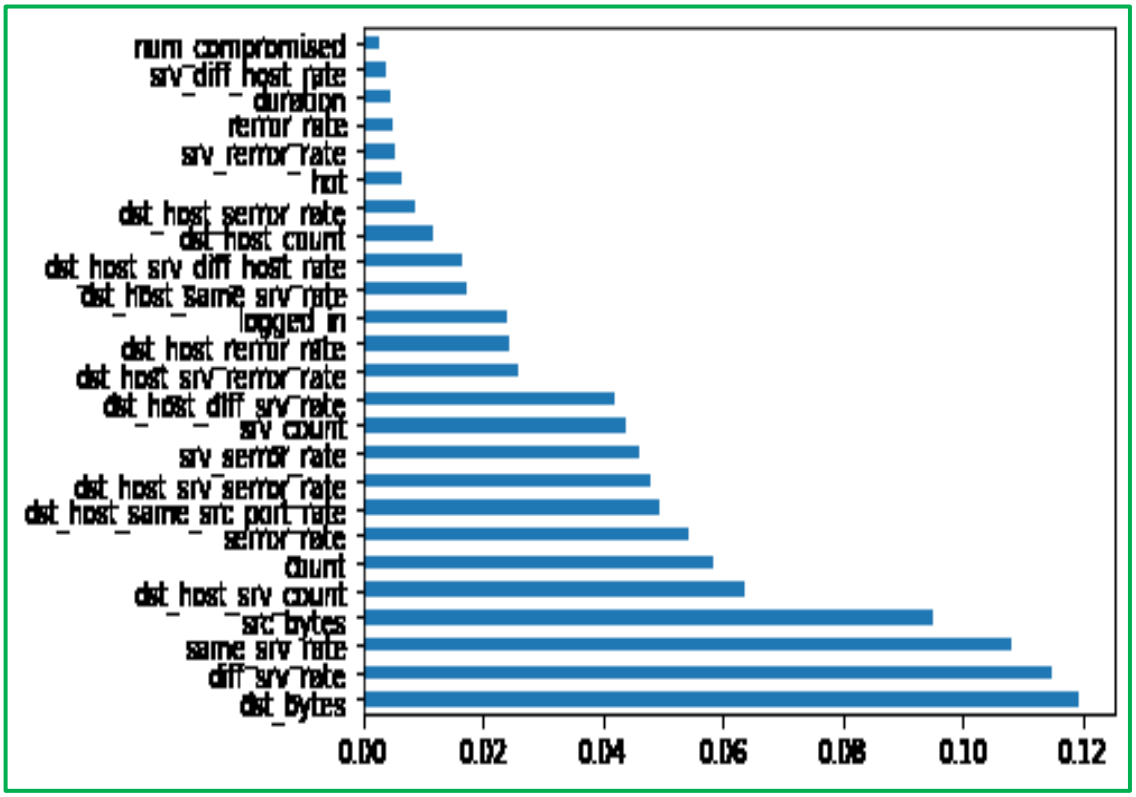


Figure 2: Important Features from RFE

IV. Build of model for classification in Keras:

A. 1.Build by Deep Neural Network for binary class

A fully-connected network structure with three layers (input, hidden and output) is used. Fully connected layers are defined using the Dense class, which can specify the number of neurons or nodes in the layer as the first argument, and specify the activation function using the activation argument. then the rectified linear unit activation function referred to as ReLU is applied on the first two layers and the Sigmoid function to ensure our network output is between 0 and 1 in the output layer. A dropout regularization is a technique used to avoid overfitting when training. Figure III shows Build Model of DNN for Binary Class.

Can piece it all together by adding each layer in DNN:

The input layer of data with 25 features since important features that select from RFE technique (the input_dim=25 argument).

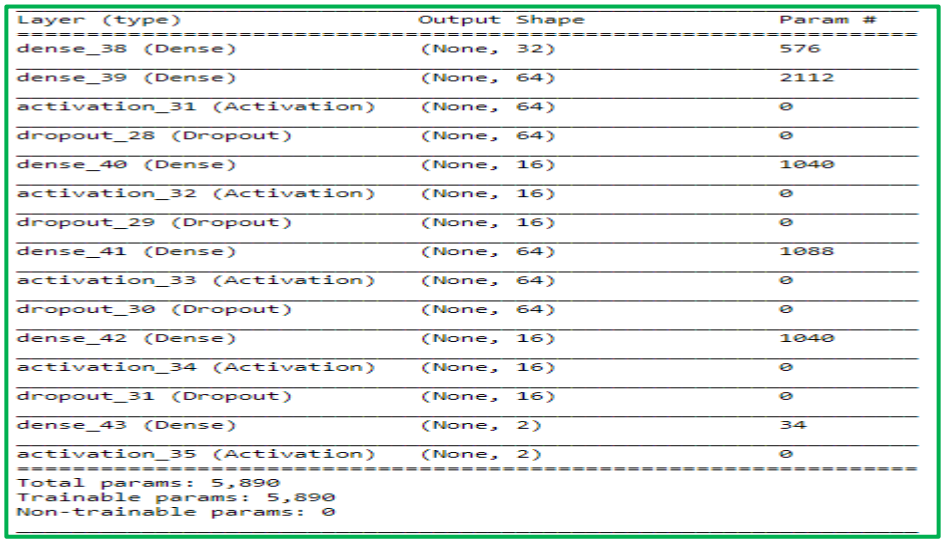
The first hidden layer has 32 nodes and uses the relu activation function.

The second hidden layer has 32 nodes and uses the relu activation function.

The third hidden layer has 16 nodes and uses the relu activation function.

The fourth hidden layer has 16 nodes and uses the relu activation function.

The output layer has Dense = 2 nodes for binary-class classification and uses the Sigmoid activation function.



Layer (type)	Output Shape	Param #
dense_38 (Dense)	(None, 32)	576
dense_39 (Dense)	(None, 64)	2112
activation_31 (Activation)	(None, 64)	0
dropout_28 (Dropout)	(None, 64)	0
dense_40 (Dense)	(None, 16)	1040
activation_32 (Activation)	(None, 16)	0
dropout_29 (Dropout)	(None, 16)	0
dense_41 (Dense)	(None, 64)	1088
activation_33 (Activation)	(None, 64)	0
dropout_30 (Dropout)	(None, 64)	0
dense_42 (Dense)	(None, 16)	1040
activation_34 (Activation)	(None, 16)	0
dropout_31 (Dropout)	(None, 16)	0
dense_43 (Dense)	(None, 2)	34
activation_35 (Activation)	(None, 2)	0

Total params: 5,890
Trainable params: 5,890
Non-trainable params: 0

Figure 3: Build Model of DNN for Binary Class

B. Build by Recurrent Neural Network for multi-class

The system uses a fully-connected network structure with three layers (input, hidden and output). input layer can specify the number of neurons or nodes in the layer as the first argument, In RNN the important argument is return sequences, by default is set to False, so return sequences=True for add more layers. Also the system use dropout regularization technique to avoid over fitting through training. Finally, in the output layer the sigmoid function is used to ensure the network output is between 0 and 1. Figure 4 shows the model of RNN for Multi-Class.

The layers can be grouped together by adding each layer in RNN:

The input layer of data with 25 variables since important features that select from RFE technique (the input_dim=25 argument)

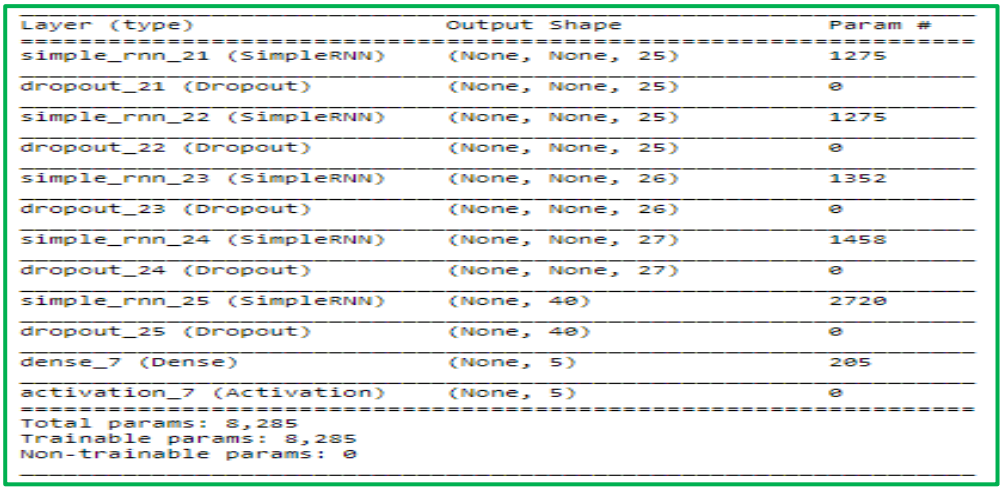
The first hidden layer has 64 nodes and uses the return_sequences=True

The second hidden layer has 64 nodes and uses the return_sequences=True

The third hidden layer has 64 nodes and uses the return_sequences=True

The fourth hidden layer has 64 nodes and uses the return_sequences=False

The output layer has Dense= 5 nodes for multi-class classification and uses the sigmoid activation function.



Layer (type)	Output Shape	Param #
simple_rnn_21 (SimpleRNN)	(None, None, 25)	1275
dropout_21 (Dropout)	(None, None, 25)	0
simple_rnn_22 (SimpleRNN)	(None, None, 25)	1275
dropout_22 (Dropout)	(None, None, 25)	0
simple_rnn_23 (SimpleRNN)	(None, None, 26)	1352
dropout_23 (Dropout)	(None, None, 26)	0
simple_rnn_24 (SimpleRNN)	(None, None, 27)	1458
dropout_24 (Dropout)	(None, None, 27)	0
simple_rnn_25 (SimpleRNN)	(None, 40)	2720
dropout_25 (Dropout)	(None, 40)	0
dense_7 (Dense)	(None, 5)	205
activation_7 (Activation)	(None, 5)	0

Total params: 8,285
Trainable params: 8,285
Non-trainable params: 0

Figure 4: Model of RNN for Multi- Class

8. RESULTS AND DISCUSSIONS

In early 2015, Keras had the first reusable open-source Python implementations for developing and evaluating deep learning models.The system was implemented by (NSL -KDD) dataset. This work was implemented by the Python language. The NSL -KDD composed of 125,973 train set and 22,544 test set confined with 5 attacks.

This section illustrates the detailed results for the proposed system and for Binary as well as Multi-class classification**.** The suggested Network Intrusion Detection System is implemented by NSL - KDD dataset and evaluated the work by Recurrent and deep neural network algorithms. The NSL - KDD is divided automatically into train and test sets. For train set, most of the DNN and RNN network topologies showed that the train accuracy is reached to 99%, The system is implemented by two algorithms for classification, deep neural network algorithm for binary class and Recurrent Neural Network algorithm for multi class. The result of classification in test set is showed as follows:-

During testing phase, results of DNN for binary-class . The result of DNN obtained best accuracy for the added 4 layers in hidden layer as compared to the other layers.

DNN results in terms of accuracy, gives good results in Binary classification for True positive rate (TPR) and false positive rate (FPR) as shown in Table III.

TABLE III: Test outcomes of DNN

Architecture	Precision	Recall	F1-score	TPR	FPR	Accuracy
DNN 4 Layer	91%	92%	77%	0.70	0.06	94%
DNN 3 Layer	89%	89%	78%	0.64	0.09	92.54%
DNN 2 Layer	81%	78%	78%	0.67	0.08	92.11%
DNN 1 Layer	82%	78%	78%	0.67	0.07	91.37%

Test results of RNNs for Multi-class classification are illustrated in table(4), the performance of RNN is evaluated on test set, for all the classes, the system was obtained good accuracy at the four added layers as compared to the other layers. RNN results in terms of accuracy, gives good results in multi classification., as shown in Table IV.

TABLE IV:Test outcomes of RNN

Architecture	Normal			Dos			Probe		
	TPR	***FPR***	***ACC***	***TPR***	***FPR***	***ACC***	***TPR***	***FPR***	***ACC***
RNN 4 Layer	0.61	0.06	0.92	0.94	0.12	0.96	0.97	0.57	0.87
RNN 3 Layer	0.63	0.70	0.92	0.94	0.08	0.96	0.97	0.57	0.87
RNN 2 Layer	0.63	0.06	0.91	0.94	0.12	0.96	0.97	0.53	0.82
RNN 1 Layer	0.67	0.07	0.91	0.94	0.11	0.96	0.97	0.55	0.81

Table IV continued

Architecture	R2L			U2R		
	TPR	***FPR***	***ACC***	***TPR***	***FPR***	***ACC***
RNN 4 Layer	0.99	0.99	0.69	100	100	0.94
RNN 3 Layer	0.99	0.99	0.80	100	100	0.92
RNN 2 Layer	100	0.99	0.69	100	100	0.91
RNN 1 Layer	100	0.99	0.75	100	100	0.89

The operating characteristics of the receiver can be used to provide explanations for the required work and work to extract the results by providing the graphics, figure 5 and figure 6 explained the results, Blue color indicates the evaluation process of training and Green indicates the evaluation process of testing process. The ROC curve for NSL-KDD. In most cases, DNN and RNN performed well a figure used as shown.

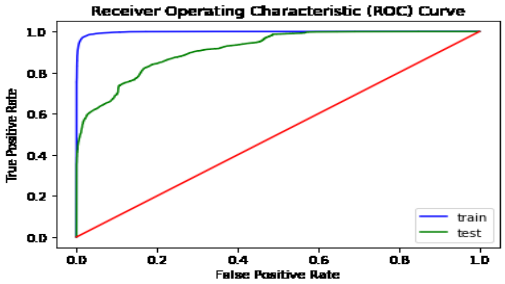


Figure 5: ROC curve for DNN

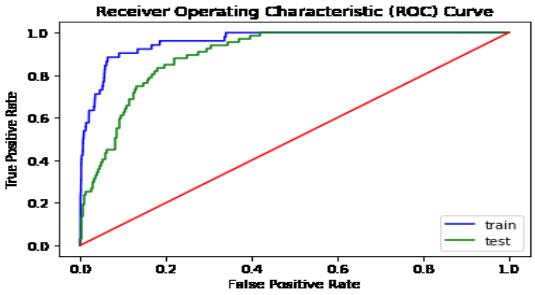


Figure 6: ROC curve for RNN

9. Conclusions and Future Work

Intrusion detection system is a technique that can be used for discovering the known and unknown intrusions before the attacker harm the devices of the networks. In this paper, a proposed Network intrusion detection alert system is implemented to build an effective and flexible Network intrusion detection system by NSL KDD dataset. By using a Recurrent Neural Network and Deep Neural Network algorithms.

At first, Deep Neural Network for binary classification is implemented and resulted accuracy rate reached 94% and False Positive Rate equal to 0.08 also, true positive rate was 92%. Then using Recurrent Neural Network for multi- class classification (DOS, Normal, probe, R2L and U2R.) with testing dataset and the accuracy is equal to 94 %. Accuracy was 96%, true positive rate 77% for Normal, accuracy was 96% and true positive rate 94% for DOS, accuracy was 87% and true positive rate 87% for Probe, accuracy was 70% and true positive rate 87% for R2L and accuracy was 94%. true positive rate (0.99) % for U2R. False alarm rate for all classes was between (0.1) and (0.8). Overall model performance was good, especially in anomaly detection.

The next of this work can be extended in 3 directions: first, it is possible to apply the system on other intrusion dataset such as Kyoto, WSN-DS and CICIDS2017. Secondly, use another feature selection technique such as LDA and rough set and other. Third, implementing the suggested system online.

References

[1] Ahmed M., Nasser Mahmood A., Hu J. (2016); A survey of network anomaly detection techniques, Journal of Network and Computer Applications, 60, 19–31, 2016.

[2] Bhuyan M. H., Bhattacharyya D. K., Kalita J. K. (2014); Network anomaly detection: methods, systems and tools, Communications Surveys & Tutorials, IEEE, 16(1), 303–336,2014.

[3] .Abd, Dhafar Hamed, and Tameem Hameed Obaida. "A Robust Approach for Mixed Technique of Data Encryption Between DES and RC4 Algorithm." Journal of Kufa for Mathematics and Computer 3.2 (2016): 48-54.

[4] Vinayakumar, R., et al. "Deep learning approach for intelligent intrusion detection system." IEEE Access 7 (2019): 41525-41550.

[5] P. Gogoi et al," Packet and flow based network intrusion dataset."Contemporary Computing". Springer Berlin Heidelberg, 2012. P 322-334.

[6] L. Li, Y. Yu, S. Bai, Y. Hou, and X. Chen, ''An effective two-step intrusion detection approach based on binary classification and k-NN,'' IEEE Access, vol. 6, pp. 12060–12073, 2018

[7] Yin, C.; Zhu, Y.; Fei, J.; He, X. A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks. IEEE Access 2017, 5, 21954–21961.

[8] I. S. Al-Mejibli, D. H. Abd, J. K. Alwan and A. J. Rabash, "Performance Evaluation of Kernels in Support Vector Machine," 2018 1st Annual International Conference on Information and Sciences (AiCIS), Fallujah, Iraq, 2018, pp. 96-101, doi: 10.1109/AiCIS.2018.00029.

[9] D. H. Abd and I. S. Al-Mejibli, "Monitoring System for Sickle Cell Disease Patients by Using Supervised Machine Learning," 2017 Second Al-Sadiq International Conference on Multidisciplinary in IT and Communication Science and Applications (AIC-MITCSA), Baghdad, Iraq, 2017, pp. 119-124, doi: 10.1109/AIC-MITCSA.2017.8723006.

[10] .Farnaaz, N.; Jabbar, M.A. Random forest modeling for network intrusion detection system. Procedia Comput. Sci. 2016, 89, 213–217.

[11] Kim, G.; Lee, S.; Kim, S. A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Syst. Appl. 2014, 41, 1690–1700.

[12] Al-Mejibli, Intisar Shadeed, Jwan K. Alwan, and Hamed Abd Dhafar. "The effect of gamma value on support vector machine performance with different kernels." International Journal of Electrical and Computer Engineering 10.5 (2020): 5497.

[13] Abd, Dhafar Hamed, Ahmed T. Sadiq, and Ayad R. Abbas. "Political Articles Categorization Based on Different Naïve Bayes Models." International Conference on Applied Computing to Support Industry: Innovation and Technology. Springer, Cham, 2019.

[14] Zaman, S.; Karray, F. Features selection for intrusion detection systems based on support vector machines. In Proceedings of the IEEE Consumer Communications and Networking Conference, Las Vegas, NV, USA, 10–13 January 2009; pp. 1–8.

[15] Kakavand, M.; Mustapha, N.; Mustapha, A.; Abdullah, M.T. Effective Dimensionality Reduction of Payload-Based Anomaly Detection in TMAD Model for HTTP Payload. KSII Trans. Internet Inf. Syst. 2016, 10, 3884–3910.

[16] Tahir, H.M.; Said, A.M.; Osman, N.H.; Zakaria, N.H.; Sabri, P.N.A.M.; Katuk, N. Oving K-means clustering using discretization technique in network intrusion detection system. In Proceedings of the 3rd International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, Malaysia, 15–17 August 2016; pp. 248–252.

[17] Vinayakumar, R.; Soman, K.P.; Poornachandran, P. Applying convolutional neural network for network intrusion detection. In Proceedings of the International Conference on Advances in Computing, Communications and Informatics (ICACCI), Udupi, India, 13–16 September 2017; pp. 1222–1228.

[18] Canbay, Yavuz, and Seref Sagiroglu. "A hybrid method for intrusion detection." 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA). IEEE, 2015.

[19] A. Javaid, Q. Niyaz, W. Sun, and M. Alam, ''A deep learning approach for network intrusion detection system,'' in Proc. 9th EAIInt. Conf.

[20] Jiang, Feng, et al. "Deep learning based multi-channel intelligent attack detection for data security." IEEE transactions on Sustainable Computing (2018).

[21] C. Yin, Y. Zhu, J. Fei, and X. He, ''A deep learning approach for intrusion detection using recurrent neural networks,'' IEEE Access, vol. 5, pp. 21954–21961, 2017

[22] NSL-KDD dataset, obtainable by: https://www.unb.ca/cic/datasets/, September 2017

[23] KDD'99 datasets, Available on: http://kdd.ics.uci.edu/databases/kddcup 99/kddcup99.html, September, 2017.

[24] T. Mikolov, M. Karafi´at, L. Burget, J. Cernock`y, and S. Khudanpur, "Recurrent neural network based language model." in Interspeech, vol. 2, 2010, p. 3

[25] Sutskever, Ilya, Oriol Vinyals, and Quoc V. Le. "Sequence to sequence learning with neural networks." Advances in neural information processing systems. 2014.

[26] Watanabe, Shinji, and Jen-Tzung Chien. Bayesian speech and language processing. Cambridge University Press, 2015.

[27] Hillman, David, and Ulrika Maude, eds. The Cambridge Companion to the Body in Literature. Cambridge University Press, 2015.

[28] Mehibs, Shawq Malik, and Soukaena Hassan Hashim. "Proposed network intrusion detection system in cloud environment based on back propagation neural network." Journal of University of Babylon for Pure and Applied Sciences 26.1 (2018): 29-40.