ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 4, April 2019

Intrusion Detection System using Recurrent Neural Network with Deep Learning

Prajyot S. Autade[1], Prakash N. Kalavadekar[2]

PG Student, Dept. of Computer Engg, Sanjivani College of Engineering, Kopargaon, Maharashtra, India[1]

Associate Professor, Dept. of Computer Engg, Sanjivani College of Engineering, Kopargaon, Maharashtra, India[2]

ABSTRACT: Due to the increasing use of the Internet, the way in which individuals are living, working and studying is changing, which causes more and more serious security threats in day-to-day life. In ensuring information security, Intrusion detection plays a significant role and it is a key technology to accurately recognize different attacks in the network. With the state of the art performance of the Deep Learning based Models in the field of computer vision, natural language processing, and speech recognition, and various Deep learning techniques are now applied to the field of cyber security. In this dissertation, a Deep learning approach has been applied for Intrusion Detection using a recurrent neural network. The critical part of developing an IDS is preprocessing of obtained network data and identifying more significant features. Hence in the design of IDS, Feature Selection method has been applied to identify a suitable subset of features. The IDS model is trained using NSL-KDD Dataset for both binary and multiclass classification. The model performance is studied by selecting a different number of features. The proposed model gives better accuracy of the intrusion detection as compare to traditional classification methods, such as J48, naive Bayesian, SVM, and random forest and provides a new research method for intrusion detection.

KEYWORDS: Intrusion detection, Recurrent Neural Network, Feature Selection, Deep Learning.

I.INTRODUCTION

Due to the improvement of information and communication techniques, an increasing amount of information is shared online. The use of the Internet in society is increasing day by day and the way in which people are living, studying and working is also changing which leads to more and more serious security threats to be faced in day-to-day life. Identifying different network attacks, especially unpredicted attacks, is an unavoidable key technical issue. An Intrusion Detection System (IDS) is a significant research achievement in the cyber security field, which can identify an attack, which could be an ongoing attack or an attack that has already occurred. Intrusion detection is similar to a classification problem, such as a binary or a multi-class classification problem. In binary classification, identify whether network traffic behavior is normal or anomalous, and in multi-class .i.e., a five-class classification problem, identify whether it is normal or any one of the other four attack types: DOS (Denial of Service), U2R (User to Root), Probe (Probing) and R2L (Root to Local). The main motivation of intrusion detection is to successfully identify the intrusive behavior and increasing the accuracy of classifiers

According to an object of observation of Jihyun Kim in [2], there are two types IDS i.e., Host-based IDS (HIDS) and Network based IDS (NIDS). The HIDS, observes the host system states and operations to detect the system event such as unauthorized access or installation. It also performs various checks such as whether there is an expected data in the state of ram or file system or not, but analyzing the behaviors related to the network is not in its scope. The NIDS observes and analyzes real-time network traffic as it is placed on the choke point of the network edge. It detects unauthorized intrusions or malicious attacks. The detection can be behavior-based intrusion detection or knowledge-based intrusion detection. The behavior-based intrusion detection is also called anomaly detection which catches attacks by comparing a normal behavior to abnormal behavior. Knowledge-based intrusion detection is also called misuse detection which detects the attacks based on previously known knowledge.

Traditional machine learning approaches have been widely used in Intrusion detection for identifying different types of attacks in the network, and the machine learning methodologies also help the network administrator to take the

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 4, April 2019

preventive measures for detected intrusions. The massive intrusion data classification problem that arises in the face of a real network application environment, cannot be effectively solved by most of the existing machine learning approaches as they belong to shallow learning and they are based on feature engineering and selection. Due to the dynamic growth of data sets, multiple classification tasks decreases accuracy. Also, shallow learning is an inappropriate method to intelligent analysis and the forecasting requirements of high-dimensional learning with a large amount of data. To create much better models, the deep learners have the great potential to extract better representations from the data, hence the rapid development is experienced by Intrusion detection technology after falling into a relatively slow period.

In 2006, the theory of deep learning was proposed by Professor Hinton[3] due to that Deep learning theory and technology has had a very rapid development. A new way to the development of the intelligent intrusion detection technology is offered by a new era of artificial intelligence which has been opened in recent years. With the growth of computational resources, there is a significant development in the domain of deep learning promoted by recurrent neural networks (RNNs). In recent years, RNNs also played a significant role in the fields of computer vision, semantic understanding, natural language processing (NLP), language modeling, picture description, speech recognition, translation, and human action recognition, among others.

In this paper, we proposed a deep learning approach for an intrusion detection system using recurrent neural networks. Intrusion detection is usually similar to a classification problem, which can be a binary or a multi-class classification problem. In binary classification, identify whether network traffic behavior is normal or anomalous and in multi-class classification i.e., five-category classification, in which identify whether it is normal or any one of the four attack types given above. The NSL-KDD dataset which covers training set and testing set with the different number of records and four types of attack is selected. The critical part of developing an IDS is preprocessing of obtained network data and identifying more significant features. Therefore in preprocessing, numericalization and normalization are applied to the dataset. Then Feature Selection method will be applied to identify a suitable subset of features. The performance of the proposed model will be evaluated in both binary classification and multiclass classification, and the impact of optimal features on the accuracy will be studied.

In the next section, we make a review of algorithms that handle high dimensional imbalanced data. In section 3 we introduce system architecture and section 4 describes system analysis. Paper is concluded in section 5.

II.LITERATURE REVIEW

In recent years, a branch of machine learning which is deep learning has become more and more popular and also it has been applied for intrusion detection; previous studies have shown that deep learning has potential to create better models than traditional methods.

Chuanlong Yin[1] explored, modeling an intrusion detection system based on deep learning, and he proposed a deep learning approach based on recurrent neural networks for intrusion detection (RNN-IDS). He studied the performance of the model on the NSL-KDD dataset, in binary classification and multiclass classification. He also analyzed the impact of the number of neurons and different learning rate on the performance of the proposed model and compared it with those of J48, artificial neural network, random forest, support vector machine, and other machine learning methods proposed by previous researchers on the benchmark data set. OzgurDepren [2] used both anomaly and misuse detection approaches and proposed a hybrid Intrusion Detection System (IDS) architecture. This new Intrusion Detection System architecture consists of two detection modules i.e., a misuse detection module, an anomaly detection module and also a decision support system for combining the results of these two detection modules. He used a Self-Organizing Map (SOM) structure for the proposed anomaly detection module to model normal behavior and the deviation from the normal behavior is classified as an attack. To classify various types of attacks, he used J.48 decision tree algorithm. The performance of the proposed hybrid IDS architecture was evaluated using the benchmark KDD Cup ’99 datasets.

The theory of deep learning was proposed in 2006 by Professor Hinton [3]. Deep learning technique allows computational models which are consist of multiple processing layers in order to learn better representations of data including multiple levels of abstraction. There was a dramatically improved the development due to these methods in

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 4, April 2019

speech recognition, object detection, visual object recognition and many other domains such as drug discovery and genomics. Deep-learning methods have multiple levels of representation, which is obtained by composing simple but non-linear modules that each transform the representation at one level (starting with the raw input) into representation at a higher i.e., more abstract level. Very complex functions can be learned with the composition of enough such transformations. Nathan Shone [4] used non-symmetric deep autoencoder (NDAE) for unsupervised feature learning in a network intrusion detection system. In the construction of the proposed model, he used stacked NDAEs.The model combines two techniques, shallow learning, and deep learning, which are capable of accurately analyzing a wide range of network traffic. The power of stacking proposed Nonsymmetric Deep Auto-Encoder (NDAE) (deep learning) was combined with the accuracy and speed of Random Forest (RF) (shallow learning). Graphics processing unit (GPU)-enabled TensorFlow has been used in the implementation of the classifier and the model was evaluated using the two benchmark datasets which are KDD Cup '99 and NSL-KDD.

Ping Wang[5] proposed a network intrusion detection system using convolutional neural networks (CNN) based classifier for enhancing the precision of model based on LeNet-5 to classify the network threat. He proposed LaNet-5 model with the adaptive delta optimization algorithm to fine-tune the model parameters and minimize a classification error by using error derivatives of back-propagation and quick response to intrusion detection using Tensor flow. To improve classification speed, he selected a set of the reduced feature using information gain (IG) scheme where the number of features selected (32 at IG > 0.119). BhupendraIngre [6] used NSL-KDD dataset to evaluate the performance of Artificial Neural Network and obtained a result for the binary class as well as five class classification (normal vs four types of attack). He used various performance measures to analyze the results and better accuracy was found. For NSLKDD dataset, the detection rate obtained was 81.2% and 79% for intrusion detection and attack type classification task respectively. When compared with the existing scheme, the higher detection rate has been achieved not only in the binary classification but also in the five class classification problems.

Howon Kim[7] used Long Short Term Memory(LSTM) architecture in a Recurrent Neural Network(RNN) which is a deep learning approach. To train the model and evaluated the performance of the model, he used KDD Cup 1999 dataset. Through the experiments, he confirmed the detection rate and false alarm rate and found an optimal hyper-parameter for LSTM-RNN. MahbodTavallaee [8] analyzed the performance of KDD CUP 99. To overcome the weakness of signature-based IDSs in detecting novel attacks, many researchers focused on anomaly detection, in the last decade, and in the evaluation of those systems, KDDCUP99 was the most widely used data set. He found two important issues highly affecting the performance of evaluated systems, after conducting a statistical analysis on this data set. Then NSL-KDD, a new data set, was proposed, which consists of selected records of the complete KDD data set and it does not suffer from any of mentioned shortcomings

Sang-Hyun Choi and Hee-Su Chae[9], suggested a new feature selection method attribute ratio that uses the attribute average of total and each class data to improve the efficiency of data mining algorithms by removing the redundant or irrelevant feature. They used NSL-KDD dataset to evaluate decision tree classifier to detect attacks based on the four attack categories: Dos, Probe, R2L, U2R. Anna Buczak [10] performed a literature survey of different machine learning (ML) and data mining (DM) methods used in cyber analytics in support of intrusion detection. He provided a short tutorial description of each ML/DM method. He identified, read, and summarized papers representing each method based on the number of citations or the relevance of an emerging method. Also, some well-known cyber data sets used in ML/DM are described as data is so important in ML/DM approaches. He addressed the complexity of ML/DM algorithms, present the discussion of challenges for using ML/DM for cyber security, and also some recommendations on when to use a given method are provided.

III. SYSTEM OVERVIEW

The recurrent neural network is shown in figure 1, which includes three different types of units such as input units, output units, and hidden units. The most important work is completed be Hidden units, and they also remember the end-to-end information i.e., hidden units are the storage of the whole network.

There is a one-way flow of information in the RNN model, from the input units to the hidden units. When we unfold the RNN, we can find that it embodies deep learning. For supervised classification learning, we can use RNNs based approach. The essential difference in Recurrent Neural Networks and traditional Feed-forward Neural Networks

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 4, April 2019

(FNNs) is that RNN has introduced a directional loop which is used memorize the previous information and so that it can be applied to the current output. The preceding output is also related to the current output of a sequence, and the nodes between the hidden layers also have connections. The output of the input layer, as well as the output of the last hidden layer, acts as the input of the hidden layer.

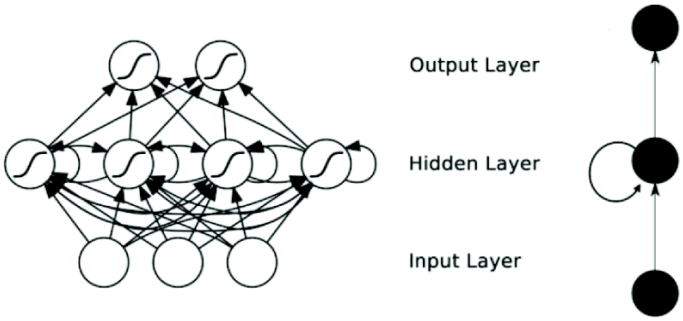


Fig. 1 Recurrent Neural Network

The architecture of the system is shown in figure 2. First, preprocessing is applied to the given input which includes numericalization and normalization. In numericalization, nonnumeric features are converted into numeric features by using encoding and in normalization, features are scaled i.e. the value of every feature is mapped to [0,1] range. In the feature selection step, optimal features are selected a given to the training of neural network. The recurrent neural network is trained using NSL-KDD is the dataset for both binary and multiclass classification. The performance of the model is evaluated using accuracy.

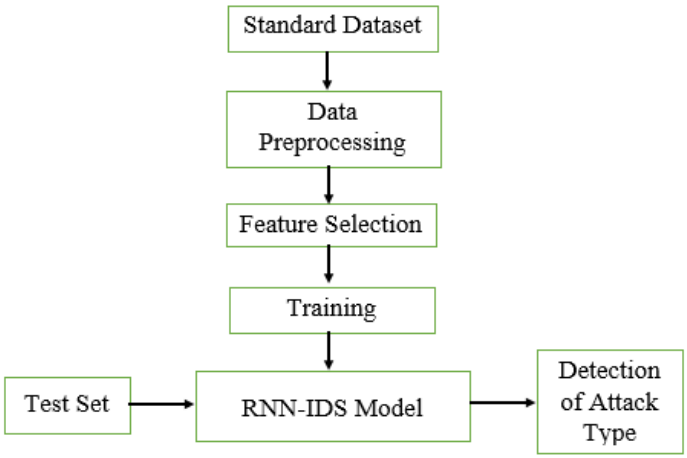


Fig. 2 Proposed System Architecture

A. Dataset Description

NSL-KDD is a dataset which covers training as well as testing sets. It is a benchmark dataset which not only effectively solves the inherent redundant records problem of the KDD cup 1999 data set but also makes the number of records reasonable in training and testing sets so that the classification does not favor more frequent records. It contains KDDTrain+ for training which has 125973 records and KDDTest+ and KDDTest-21 for testing which has 22544 and 11850 records respectively. NSL-KDD dataset has normal records and records for four different types of attacks so it can be used for binary as well as multiclass classification. The four types of attacks in NSLKDD are Denial of service attacks, Root to Local attacks, User to Root attacks and Probing attacks. There are 41 features and 1 class label for each

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 4, April 2019

record and the features includes basic features, content features, and traffic features. Some specific attack types that disappear in the training set are added to the testing set to provide a more realistic theoretical basis for intrusion detection.

B. Data Preprocessing

Numericalization: NSL-KDD dataset contains 38 numeric features and 3 non-numeric features. Recurrent neural network model requires a numeric matrix as an input, so we need to convert these non-numeric features into numeric form. The non-numeric features are 'protocol type ', 'service' and 'flag' features. For example, the feature 'protocol type 'has three types of values, 'udp', 'tcp ', and 'icmp ', and its numeric values are encoded as binary vectors (1,0,0), (0,1,0) and (0,0,1). Similarly, the feature 'service 'has 70 different types of values, and the feature 'flag 'has 11 different types of values. Continuing in this way, 41-dimensional features map into 122-dimensional features after transformation.

Normalization: For some features, the difference between the maximum and minimum values has a very large scope. Such features are 'duration[0,58329] ', 'src bytes [0,1.3X109] 'and 'dst bytes [0,1.3X109] '. Apply the logarithmic scaling method for scaling to these features to obtain the ranges of 'duration[0,4.77]', 'src bytes [0,9.11] 'and 'dst bytes [0,9.11]'. Then map every feature to the [0, 1] range linearly using equation (1), where Max is the maximum value and Min is a minimum value for each feature.

$$x_i = \frac{x_i - Min}{Max - Min} \quad (1)$$

C. Feature Selection

Feature selection is a process of selecting a subset of features using certain criteria. Most of the time, data includes noisy, redundant and irrelevant features, which should be removed using feature selection. Feature selection reduces both number of features as well as computational complexity of the algorithm. It also increases the speed of algorithm and improves learning accuracy. NSL-KDD dataset has 41 features with one class attribute. But not all features are important, few features have no role or minimum role in the intrusion detection system. There are various feature selection methods such as information gain(IG), the gain ratio (GA), correlation-based feature selection(CFS) and Attribute ratio(AR), which is calculated by average or frequency of features. In this work two methods are used .i.e. Information Gain and Attribute Ratio for selecting optimum features.

D. Training

Forward Propagation: Forward propagation is the step to do after initializing model, to check its performance i.e. how well the neural network is behaving. Forward propagation in neural networks is to predict the output values and compare it with the real/actual value to get the error or loss. The actual value and predicted values are used to calculate the loss. As the flow of calculation is going in the natural forward direction i.e., from the input -> through the neural network-> to the output hence it is called forward-propagation. In forward propagation, each layer has an activation function.

Weights Update: Weight update is done by using backpropagation which means backward propagation of errors. In this first, compute the partial derivatives of loss with respect to weight matrices and bias vectors, then propagate backward to update the weights.

E. Testing

Testing is the final step where model performance is evaluated. The model performance is predicted using KDDTest+ and KDDTest-21 sets which are present in NSL-KDD dataset. Confusion matrix can be used to calculate the accuracy of the model which is the most important performance indicator used to measure the performance of the RNN model.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 4, April 2019

IV. SYSTEM ANALYSIS

A. Evaluation Metrics

Accuracy is the most important performance indicator of intrusion detection used to measure the performance of the RNN model. The confusion matrix is used to find the accuracy which is shown in figure 3. It contains TP, FP, TN, FN.

- True Positive (TP) means correctly rejected i.e., the number of anomaly records that are identified as anomaly.
- False Positive (FP) means incorrectly rejected i.e., number of normal records that are identified as anomaly.
- True Negative (TN) means correctly admitted i.e., the number of normal records that are identified as normal.
- False Negative (FN) means incorrectly admitted i.e., number of anomaly records that are identified as normal.

Predicted Class / Actual Class	anomaly	normal
anomaly	TP	FN
normal	FP	TN

Fig. 3 Confusion matrix.

Accuracy is the percentage of the number of records classified correctly to the total of records shown in equation(2).

$$AC = \frac{TP+TN}{TP+TN+FP+FN} \qquad (2)$$

B. Results

In this, we have used one of the open source neural network library written in Python. The experiment is performed on a personal laptop, which has a configuration of AMD E1-6010 APU with AMD Radeon R2 Graphics @ 1.35 GHz, 4 GB memory. The number of epochs is given 100, number of hidden nodes is 80 and learning rate is 0.01.

1) Binary Classification

For binary classification without feature selection, 41 features are used and for binary classification with feature selection, we observe the classification accuracy on the NSL-KDD dataset with different number of features as shown in table1. In our experiment, the model gets a higher accuracy, when the number of features selected are 24.

Table1: The accuracy of RNN with different number of features for binary classification

No. of features	Information Gain(IG)			Attribute Ration(AR)		
	KDDTrain+	KDDTest+	KDDTest-21	KDDTrain+	KDDTest+	KDDTest-21
16	99.73	77.72	57.73	98.01	74.71	55.72
17	99.71	80.05	62.61	97.96	75.68	57.29
18	99.73	78.76	60.12	98.02	71.40	58.10
19	99.75	80.39	62.90	98.02	75.58	60.81
20	99.78	81.85	65.83	98.48	77.50	58.90
21	99.78	81.29	64.57	98.94	74.84	55.46
22	99.76	79.79	61.71	98.97	76.53	60.25
23	99.78	80.46	62.91	99.41	76.87	58.98
24	**99.81**	**82.83**	**67.77**	**99.38**	**79.28**	**63.10**

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 4, April 2019

25	99.89	79.89	61.98	99.17	75.94	56.68
26	99.86	79.52	61.33	99.37	77.93	63.71
27	99.85	81.69	66.09	99.32	68.29	64.05
28	99.86	82.82	67.37	99.48	75.71	60.14
29	99.86	80.57	63.13	99.43	75.59	57.23

The experimental results show that, when features selection is not used, the model gives 81.27% and 64.43% accuracy for KDDTest+ and KDD Test-21 dataset respectively. Using feature selection(IG), model gives 82.83% and 67.77% accuracy for KDDTest+ and KDD Test-21 dataset and using feature selection(AR), model gives 79.28% and 63.10% accuracy for KDDTest+ and KDD Test-21 dataset respectively.

2) Multiclass Classification

For Multiclass classification, the classification accuracy on the NSL-KDD dataset with different number of features is shown in table2.

Table2: The accuracy of RNN with different number of features for Multiclass classification

No. of features	Information Gain(IG)			Attribute Ration(AR)		
	KDDTrain+	KDDTest+	KDDTest-21	KDDTrain+	KDDTest+	KDDTest-21
16	99.76	26.21	48.54	98.01	76.09	57.72
17	99.76	30.13	48.69	98.45	73.69	49.99
18	99.78	30.98	45.31	98.43	72.96	48.75
19	99.78	33.78	45.09	**98.68**	**76.12**	**54.70**
20	99.79	33.45	49.76	99.49	22.94	32.62
21	99.82	29.08	43.13	99.54	52.79	37.44
22	99.82	32.81	48.46	99.57	34.80	37.29
23	99.81	39.43	52.21	99.65	34.22	37.82
24	99.82	43.90	57.21	99.75	42.68	42.05
25	99.85	40.29	55.91	99.71	28.98	32.75
26	99.86	37.89	52.10	99.75	41.82	48.26
27	**99.87**	**55.74**	**47.34**	99.77	38.71	45.47
28	99.86	42.92	49.34	99.78	38.76	44.18
29	99.87	37.28	50.81	99.78	38.35	44.83

The experimental results show that, when features selection is not used, the model gives 79.77% and 61.85% accuracy for KDDTest+ and KDD Test-21 dataset respectively. Using feature selection(IG), model gives 55.74% and 47.18% accuracy for KDDTest+ and KDD Test-21 dataset and using feature selection(AR), model gives 76.12% and 54.70% accuracy for KDDTest+ and KDD Test-21 dataset respectively.

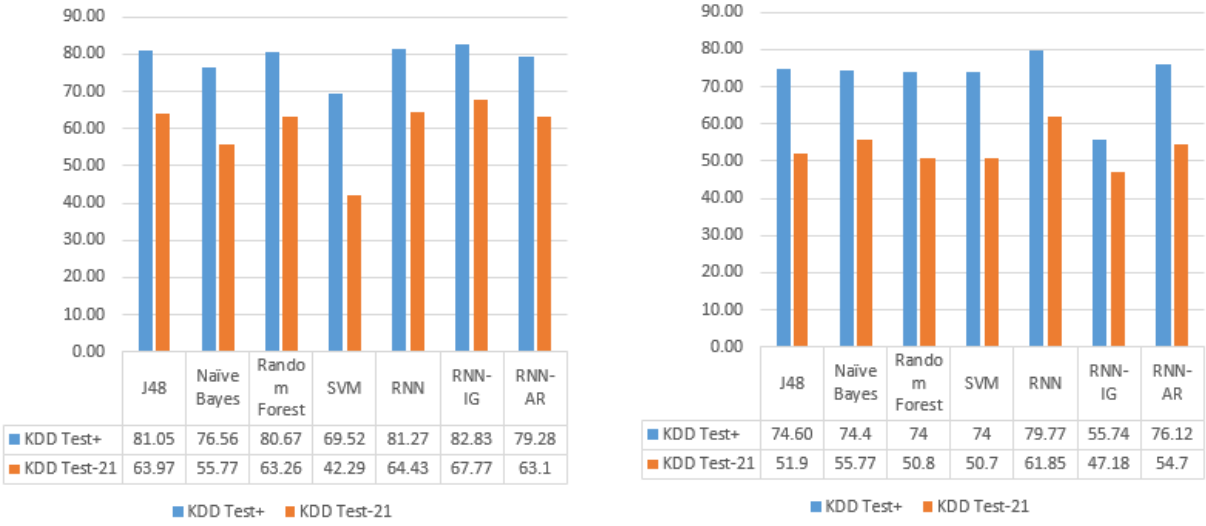
ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 4, April 2019



(a) Binary Classification (b) Multiclass Classification

Fig 4. Comparing Results of RNN with different Techniques.

In [1][8] authors have shown the results obtained for J48, Naïve Bayes, random forest and SVM. These results are all based on the same benchmark – NSL KDD dataset. The performance of the proposed RNN model is superior to other classification algorithms. In binary classification, with feature selection (IG) the model accuracy is increased as shown in figure 4(a) and in multiclass classification, with feature selection (AR) the model accuracy is still greater than other machine learning algorithms as shown in figure 4(b).

V .CONCLUSION

The Intrusion Detection System using Recurrent Neural Network model has a strong modeling ability for intrusion detection, also it has high accuracy also it has high accuracy in both binary and multiclass classification when all features are used. In binary classification, the model accuracy is increased when 24 features are selected using information gain feature selection method as compared to Attribute Ratio. In multiclass classification, model gives better accuracy when features are selected using Attribute Ratio. The proposed system uses NSL-KDD dataset to train and test model for both binary and multiclass classification. Compared with traditional classification methods, such as J48, naive Bayesian, SVM, and random forest, the performance gives a higher accuracy. The model can effectively improve both the accuracy of intrusion detection and the ability to recognize the intrusion type.

REFERENCES

[1] Chuanlong Yin, Yuefei Zhu, JinlongFei, And Xinzheng He, “A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks", vol. 05, pp. 21954-21961, oct 2017.
[2] Depren, Ozgur, et al., "An intelligent intrusion detection system (IDS) for anomaly and misuse detection in computer networks", Expert systems with Applications 29.4, pp. 713-722, 2005
[3] Y. LeCun, Y. Bengio, and G. Hinton, " Deep learning", Nature, vol. 521, no. 7553, pp. 436-444, May 2015.
[4] Nathan Shone , Tran Nguyen Ngoc, Vu DinhPhai , and Qi Shi, "A Deep Learning Approach to Network Intrusion Detection”, vol. 2, pp. 41-50 no. 1, feb 2018
[5] Wen-Hui Lin1, Hsiao-Chung Lin, Ping Wang, Bao-Hua Wu, Jeng-Ying Tsai, "Using Convolutional Neural Networks to Network Intrusion Detection for Cyber Threats” , ISBN 978-1-5386-4342-6 ,pp.1107-1170, 2018
[6] B. Ingre and A. Yadav, "Performance analysis of NSL-KDD dataset using ANN”, in Proc. Int. Conf. Signal Process. Commun. Eng. Syst., Jan. 2015, pp. 92-96.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 4, April 2019

[7] Jihyun Kim, Jaehyun Kim, Huong Le Thi Thu, andHowon Kim, "Long Short Term Memory Recurrent Neural Network Classifier for Intrusion Detection", Platform Technology and Service (PlatCon), 2016 International Conference on. IEEE, 2016.

[8] M. Tavallaee, E. Bagheri,W. Lu, and A. A. A. Ghorbani, "A detailed analysis of the KDDCUP 99 data set", in Proc. IEEE Symp. Comput. Intell. Secur. Defense Appl., Jul. 2009, pp. 1-6.

[9] Sang-Hyun Choi and Hee-Su Chae, "Feature Selection using Attribute Ratio in NSL-KDD data", International Conference Data Mining, Feb 4-5, 2014 Bali (Indonesia),pp. 90-92

[10] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection", IEEE Commun. Surveys Tuts, vol. 18, no. 2, pp. 11531176, 2nd Quart, 2016