TECHNIUM

Technium Vol. 4, No. 8 pp.64-71 (2022)
ISSN: 2668-778X
www.techniumscience.com

Detecting network attacks Model based on a long short-term memory (LSTM)

Teba Ali Jasim Ali [1], Muna M. Taher Jawhar [2]

[1] College of Computer Sciences and Mathematics, University of Mosul, Iraq
tebaa.20csp6@student.uomosul.edu.iq

[2] College of Computer Sciences and Mathematics, University of Mosul, Iraq
dr.muna_taher@uomosul. edu.iq

Abstract. Nowadays, network-connected devices such as mobile phones and IoT devices are increasing, the types and numbers of these devices are increasing, the impact of successful attacks is increasing and the fear is growing due to the security effects when using them. In addition, a broader attack surface is available to identify and respond to these network attacks, different systems are used to prevent and stop Some of these systems consist of two layers, the first layer which provides Security and Intrusion Prevention is the firewall, while the second layer is the network intrusion detection system or attack detection system, if only the first layer represented by the firewall is used we cannot prevent attack, that's why attack detection or malware detection systems are used along with a firewall.

Keywords. Network security, Deep Learning, IDS, long short-term memory.

1. Introduction

Nowadays, the demand for the Internet and Internet-related devices, as well as the data generated by these devices, has increased. Implementation of many activities through the network, so the network is one of the essential parts of life [1]. This increases the incentive to attack networks and information systems, and thus, it becomes more difficult for detection tools. The normal routine handles these attacks efficiently [2]. Since the advent of modern computing and the Internet, we have become more and more dependent on using computers to store and process data as it is transmitted over the Internet and stored in such remote storages. This rapid inflation has attracted a large number of cybercriminals to launch cyber-attacks all over the world with every day. As it passes, these attacks have become more widespread and more complex in nature [3].

That is why attack detection and intrusion detection systems (IDS) are necessary and are an effective and lightweight system that detects attacks on the network where It works to classify many electronic attacks as well as monitor events that occur in a computer system or network. Analyze and dynamically detect potential attacks and prevent undesirable access, mostly by collecting data. Automatically from a set of network sources and then analyze the data in search of security flaws [3][4].

Because of the rapid development of networks and the generation of a lot of data that requires accurate authentication and security (network security restrictions) to protect it, so Intrusion detection systems based on artificial intelligence techniques are among the most important techniques used to address security threats. Cyber security and security [5] [6] Network-based Intrusion Detection System is a security solution that protects against internal and external attacks, as well as unwanted access to the network using software and/or hardware. The most common method is a firewall, which is supposed to protect the entire network from illegal access via IP address and port number, with an attack detection

system taking over management that limits the amount of network intrusion attempts, such as Denial of Service attacks, which compromise the security of a single device or an entire network [7].

The aim of the research is to design and implement a model that classifies attacks by analyzing network traffic and based on a dataset (cic-ids2018) and algorithm (LSTM).

In the paper [8] Yin and colleagues used the RNN DL algorithm, the inputs to the network consisted of 122 neurons that when tested on the KD-NSL dataset, their model achieved an accuracy score of 83.28% in binary classification and 81.29% in multiple classification .

The researcher in [9] proposed an intrusion detection model based on RNN-type neural networks. The model built on forward propagation and reverse propagation was used with NSL-KDD data, and the detection results were for binary classification. 83.28% while for multi-layer classification it was 81.29%. Furthermore, we may be able to increase performance in the future through the use of these advanced models.

J. Kim et al. [10] developed, a CNN-based intrusion detection system, used the CIC-2018 dataset, converted numerical data into images, and then organized the convolutional layers and CNN max aggregation layers.

In [11] the proposed system implemented a dynamic system for detecting anomalies in the network using a deep learning method, long-term memory (LSTM). It also added the Attention Mechanism (AM) to improve the performance of the model while the SMOTE algorithm improved the loss function to solve the category imbalance problem in the CSE-CIC-IDS2018 dataset.

In this paper [12], the researchers built a model based on deep learning to detect cyber-attacks based on the CIC-2018 database, and the detection rate was good, nearly 90%.

In the research [13] Abdel Hamid et al. Use the automated encoder in addition to the principle of analyzing the characteristics and components of the CICIDS2017 dataset to reduce the number of features or dimensions of the resulting data, and then use characteristics from both techniques to build a classification model to detect various malicious attacks.

In our research, based on the LSTM network with reducing characteristics and binary classification, a detection accuracy of 95% and a loss ratio of 0.14% was obtained..

2. BACKGROUND THEORY

2.1. The dataset

It is a common dataset with many published works explored in the field of intrusion detection.Developed by the Amazon / AWS Web Services (AWS) development platform, this dataset provides several attack profiles that can be used in the field of intelligent security and apply them to network topologies and protocols in a general approach [14]. This dataset has been optimized taking into account the CSECIC IDS2017 standards. Data (CSE-CIC-IDS2018) contains approximately 16,000,000 samples collected over ten different days [15]. (CSE-CIC-IDS2018) was established in cooperation between the Communications Security Corporation (CSE) and the Canadian Institute of Cyber Security (CIC) in view of To change network behaviors and evolve attacks, it has become necessary to move away from static data sets and move Towards data that can be modified, expanded, and iterated [16].CSE-CIC IDS2018 is a public dataset in use today that contains 2 categorized profiles and consists of 5 different attack methods. Various data scenarios were collected, and crude was edited daily.

80 statistical properties such as packet length, number of packets, number of bytes were calculated and these features provide forward and backward directions of network flow, packets, etc. during data generation. Finally, the data set was published via an online file to all researchers. The dataset is published in CSV and PCAP formats with approximately 5 million records. CSV format is mainly used in the field of AI, and PCAP format is used to extract new features[17]. as shown in the figure The process of extracting CSV files from PCAP and then preprocessing the data before training it on the network as shown In Figure (1).

Technium Vol. 4, No. 8 pp.64-71 (2022)
ISSN: 2668-778X
www.techniumscience.com

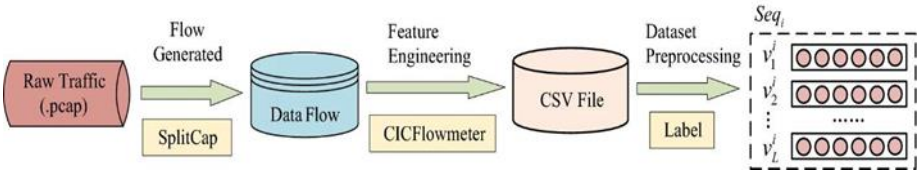

Figure (1) Extracting the network connection data set and doing a preprocess

A dataset (CSE-CIC IDS 2018) contains network traffic and system logs. These logs consist of 10 days of data subsets collected on different days by injecting 16 types of attacks. CSE-CIC IDS 2018 has a size of more than 400 GB [18].

2.2. long short-term memory (LSTM)

Long-term memory is a recurrent artificial neural network (RNN) used in deep learning. LSTM aims to overcome the short-term memory problem observed in RNN by predicting a complete sequence of structured data while maintaining a certain level of memory with respect to previous operations [19]. In contrast to normal feed-forward neural networks, LSTM allows previous outputs to be used as inputs to subsequent layers with hidden states at random intervals. In this way, we have the possibility to use the input of any length; Therefore, it is very suitable for classifying, processing, and predicting time series, such as traffic collected in a computer network, with time intervals of unknown length [20]. A long-term memory architecture is a string-based architecture, as shown In Figure(2), the information is computed in "cells", while memory is managed by "gates", and is generally classified into the following types: forget gate, input gate, and output gate. These gates have a sigmoid activation function to compress the data received by the cells and learn the necessary data to keep or forget. Finally, predictions are made by passing the relevant information to the sequence of sequences in the neural network. These networks are widely used and very useful today, being at the forefront of linguistic modeling, such as translation or text creation, in short, any activity related to reading and writing, due to their ability to recognize patterns over time.

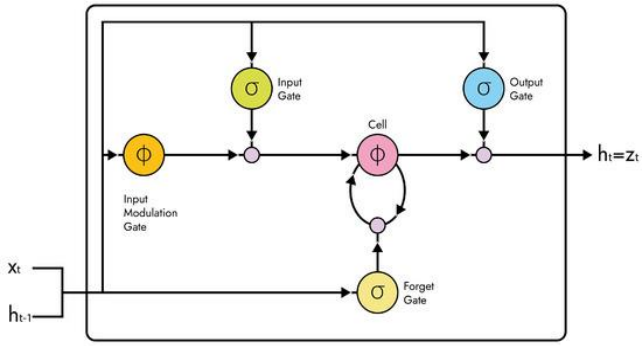

Figure (2) memory. The geometry of lstm

3. Methodology

After we downloaded a set of sub-files consisting of ten sub-records of type (.csv), nine files contain 79 properties, and the remaining file consists of 83 properties. Due to the huge size of the data set of 400 GB, each file contains two or three types of attacks, and the files in the data set can be combined to include both attack labels for processing, but files of each attack type are combined. Leads to an increase

in the size of the data set which leads to an increase in the processing time, so I created a new multi-class data set with 8 attacks in addition to the benign type (Bot, FTP Brute-force ,DoS attack-GoldenEye, Infilteration, SSH Bruteforce, DDOS attack-HOIC, and DDOS attack-LOIC-UDP) from individual datasets and randomly selected record for all CSE-CIC-IDS2018 datasets (Fri 03-02-2018, Fri 16-02-2018, Friday-23-02-2018, Thursday 15-02-2018, Thursday 03-1-2018, Wednesday 14-02-2018 Wed 21-02-2018 (CSE-CIC-IDS2018 Data Sets).

After the process of merging in one record, we got the following types, as shown in the figure*:*

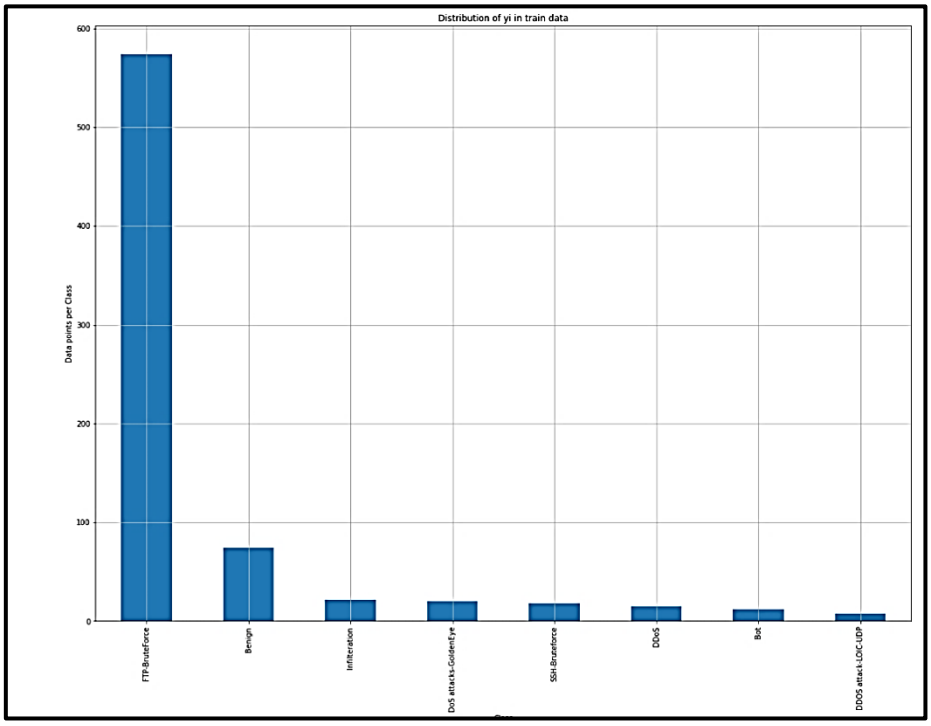

Figure (3) Types of Attacks in the CIC-2018 Dataset)

Simple Binary Analysis: As shown in Figure (5) the binary classification graph, the data set is highly skewed towards benign threats.

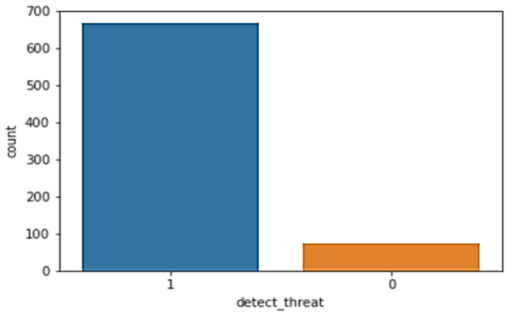

Figure (5) Binary classification of the (CIC-2018) sub-dataset.

To further clarify the data analysis, a statistical table was made for the data set as shown in Table (1), which shows the percentages and numbers of each type of attack, in addition to the benign type when analyzing our extracted data set.

Table 1. distribution of attack types for data (CIC-2018)

Label	count	percentage
FTP-BruteForce	574	0.772544
Benign	74	0.099596
Infilteration	22	0.02961
DoS attacks-GoldenEye	20	0.026918
SSH-Bruteforce	18	0.024226
DDoS	15	0.020188
Bot	12	0.016151
DDOS attack-LOIC-UDP	8	0.010767

After the statistical and visual analysis of the data, we pre-processed the data set to be suitable for applying machine learning techniques according to the following steps:

1) We process missing and infinity data in two ways. The first method removes all missing and infinite values The second method replaces the data set with infinite values with the maximum value and the missing values with the average values.

2) Feature selection is a way of selecting some features from the data and ignoring irrelevant features, so we removed some of these categorical features that lead to increased training time of the network. I also implemented feature selection (PCA) technology to transform data from a high-dimensional space to a low-dimensional space so that the low-dimensional representation retains some useful properties of the original data[4]

3) One-Hot-encoding: Features are digitized with one hot encoding that is used to convert all categorical features to binary features. The input to this must be an array of integers that will be a Sparks array with each column containing a value for one feature.

4) Categorical data encryption. Label encoder: Features are converted from categorical to scalar using the Label-encoder approach in machine learning using Python. Features are converted using a Label-encoder from a class to a number.

5) The dataset is then divided for each attack class and each attack is renamed to 0 for normal connection, and 1 for attack class (abnormal connection).

6) Split the dataset into a training set and a test set.

4. Results and discussion

A long-term memory neural network model was built as shown in Table (2) showing the parameters used, the layers used, and the input and output in each layer.

Technium Vol. 4, No. 8 pp.64-71 (2022)
ISSN: 2668-778X
www.techniumscience.com

Table 2. Layers and parameters in the network model of long-term memory on data (CIC-2018).

Layer (type)	Output Shape	Param #
lstm_2 (LSTM)	(None, 70, 8)	320
dropout_2 (Dropout)	(None, 70, 8)	0
lstm_3 (LSTM)	(None, 8)	544
dropout_3 (Dropout)	(None, 8)	0
dense_1 (Dense)	(None, 1)	9
activation_1 (Activation)	(None, 1)	0

Total params: 873
Trainable params: 873
Non-trainable params: 0

The performance of the long-term memory (LSTM) network model on (CIC-2018) data.

In the binary classification of the long-term memory network(LSTM) model, the accuracy of the model was calculated for both the training and validation samples in each epoch in which the model is trained. Figure (6) shows the accuracy of the model, as well as the calculation of the loss and the verification of the loss as in Figure (7).

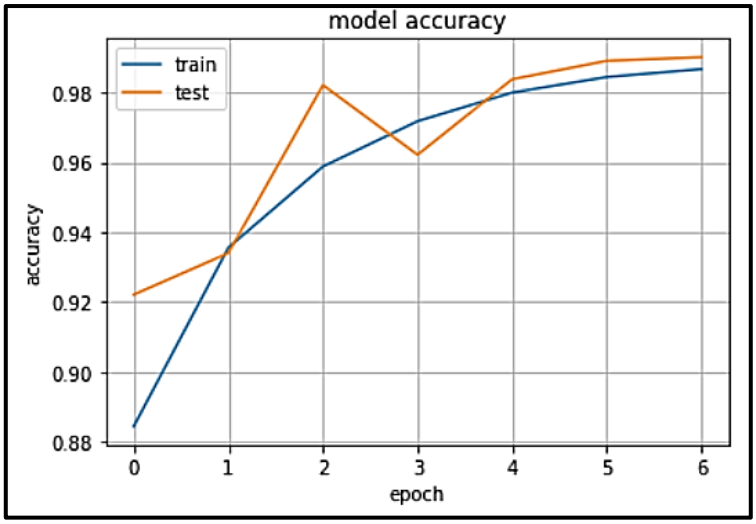

Figure (6) Model Accuracy

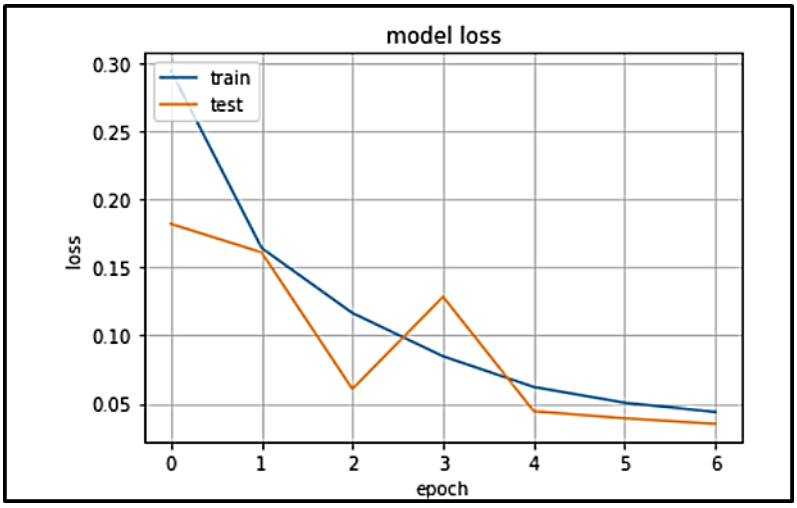

Figure (7) Loss of the Model

5. Conclusion

The goal of the project was to build a network attack detection system using deep learning techniques Communication (network traffic) is classified as "benign" or "attack", a binary classification model in which we used Network(LSTM) to classify types of connections and attacks. The system was trained and tested using a dataset (CIC-2018), and classifier accuracy, detection rates and error rates were measured using the PYTHON language and Jupiter Notebooks code editor. The results of empirical experiments show better performance when using the low-trait dataset than when using the full-trait dataset. We reached an accuracy rate (98.6%) in distinguishing between offensive and normal contact.

References

[1] S. N. Nguyen, V. Q. Nguyen, J. Choi, and K. Kim, “Design and implementation of intrusion detection system using convolutional neural network for DoS detection,” ACM International Conference Proceeding Series. pp. 34–38, 2018. doi: 10.1145/3184066.3184089.

[2] B. N. 6ORCID andWilliam J. B. 1ORCID by Andrew Churcher 1ORCID, Rehmat Ullah 2,*ORCID, Jawad Ahmad 1ORCID, Sadaqat ur Rehman 3ORCID, Fawad Masood 4, Mandar Gogate 1, Fehaid Alqahtani 5ORCID, “Sensors _ Free Full-Text _ An Experimental Analysis of Attack Classification Using Machine Learning in IoT Networks.pdf.” p. 32, 2021.

[3] A. Khurshid and G. A. Khan, “Online Machine Learning-based Framework for Network Intrusion Detection,” 2018.

[4] Y. Tang, L. Gu, and L. Wang, “Deep stacking network for intrusion detection,” Sensors, vol. 22, no. 1, 2022, doi: 10.3390/s22010025.

[5] M. Abdullahi et al., “Detecting Cybersecurity Attacks in Internet of Things Using Artificial Intelligence Methods: A Systematic Literature Review,” Electron., vol. 11, no. 2, pp. 1–27, 2022, doi: 10.3390/electronics11020198.

[6] A. Albayati, N. F. Abdullah, A. Abu-Samah, A. H. Mutlag, and R. Nordin, “A Serverless Advanced Metering Infrastructure Based on Fog-Edge Computing for a Smart Grid: A Comparison Study for Energy Sector in Iraq,” Energies, vol. 13, no. 20, 2020, doi: 10.3390/en13205460.

[7] A. Janagam and S. Hossen, “Analysis of network intrusion detection system with machine learning algorithms (deep reinforcement learning algorithm).” 2018.

[8] Z. Cui, F. Xue, X. Cai, Y. Cao, G. G. Wang, and J. Chen, "Detection of Malicious Code Variants Based on Deep Learning," IEEE Transactions on Industrial Informatics, vol. 14, no. 7. pp. 3187–3196, 2018. doi: 10.1109/TII.2018.2822680.
[9] 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. doi: 10.1109/ACCESS.2017.2762418.
[10] J. Kim, Y. Shin, and E. Choi, "An Intrusion Detection Model based on a Convolutional Neural Network," Journal of Multimedia Information System, vol. 6, no. 4. pp. 165–172, 2019. doi: 10.33851/jmis.2019.6.4.165.
[11] P. Lin, K. Ye, and C.-Z. Xu, "Dynamic network anomaly detection system by using deep learning techniques," in International conference on cloud computing, 2019, pp. 161–176.
[12] R. I. Farhan1, A. T. Maolood2, and NidaaFlaih Hassan3, "Performance analysis of flow-based attacks detection on CSE-CIC-IDS2018 dataset using deep learning." Indones. J. Electr. Eng. Comput. Sci., vol. 20, pp, p. 6, 2020. doi: doi: 10.11591/ijeecs.v20.i3.pp1413-1418.
[13] R. Abdulhammed, H. Musafer, A. Alessa, M. Faezipour, and A. Abuzneid, "Features dimensionality reduction approaches for machine learning based network intrusion detection," Electron., vol. 8, no. 3, 2019, doi: 10.3390/electronics8030322.
[14] L. D'hooge, T. Wauters, B. Volckaert, and F. De Turck, "Classification hardness for supervised learners on 20 years of intrusion detection data," IEEE Access, vol. 7, pp. 167455–167469, 2019.
[15] A. Hafid, S. Benouar, M. Kedir-Talha, F. Abtahi, M. Attari, and F. Seoane, "Full impedance cardiography measurement device using raspberry PI3 and system-on-chip biomedical instrumentation solutions," IEEE J. Biomed. Heal. informatics, vol. 22, no. 6, pp. 1883–1894, 2017.
[16] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization.," ICISSp, vol. 1, pp. 108–116, 2018.
[17] V. Jaganathan, P. Cherurveettil, and P. Muthu Sivashanmugam, "Using a prediction model to manage cyber security threats," Sci. World J., vol. 2015, 2015.
[18] D. Ravikumar, Towards Enhancement of Machine Learning Techniques Using CSE-CIC-IDS2018 Cybersecurity Dataset. Rochester Institute of Technology, 2021.
[19] A. H. Mirza and S. Cosan, "Computer network intrusion detection using sequential LSTM neural networks autoencoders," in 2018 26th signal processing and communications applications conference (SIU), 2018, pp. 1–4.
[20] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," Phys. D Nonlinear Phenom., vol. 404, p. 132306, 2020.

Submitted: 15.07.2022
Revised: 16.08.2022
Accepted: 26.08.2022