INTERNATIONAL CONFERENCE ON ADVANCES IN TECHNOLOGY AND COMPUTING (ICATC-2021)
Faculty of Computing and Technology (FCT), University of Kelaniya, Sri Lanka
18th December 2021

Index N0: CT-08-51

A Deep Neural Network Approach for Analysis of Firewall Log Data

Chandesh Lillmond (1st Author)
School of Innovative Technologies and Engineering (SITE)
University of Technology Mauritius (UTM)
Pointe-aux-Sables, Republic of Mauritius
lillmond@umail.utm.ac.mu

Geerish Suddul (2nd Author)
School of Innovative Technologies and Engineering (SITE)
University of Technology Mauritius (UTM)
Pointe-aux-Sables, Republic of Mauritius
g.suddul@umail.utm.ac.mu

***Abstract* — In this paper, we propose an intelligent approach for the classification of incoming and outgoing firewall traffic packets. A firewall is a quintessential tool that ensures the control of traffic over machines' communication over a network. It uses a set of specific rules to define the traffic and thus assists in avoiding cyber-attacks which can be very costly to an organization. Our intelligent approach is mainly through the application of the Deep Neural Network (DNN) Machine Learning algorithm so that packets going through the firewall can be automatically classified as either allow, deny or drop. Our experiments demonstrate a classification accuracy of around 94%, which is higher when compared with other approaches.**

Keywords — Firewall, DNN, Classification

I. INTRODUCTION

A firewall is software that is used to detect malicious data from incoming and outgoing packets on the internet. A firewall normally sets up a wall between a trusted network and an untrusted network, like the Internet It is important to have a firewall installed on a computer as it helps to protect the user from cyber-attacks and corrupted data. The main job of a firewall is to inspect packets transferred between computers. The firewall has a set of rules which indicate the appropriate action to each packet. There are 3 main actions regarding a packet is "allow", "deny" and "drop". With the help of machine learning, the firewall can become more accurate to classify the packets to their corresponding set of rules, which can result in a more secure network everywhere.

The main objective of this research is (i) to propose the modelling and implementation of a classifier model based on Deep Learning techniques and, (ii) to perform a comparative study that evaluates the performance of the proposed model.

II. LITERATURE REVIEW

The following segments portray different work completed by researchers utilizing different machine learning algorithms and methodologies carried out. This study [2] analysed data of a network using supervised machine learning techniques. To classify the data set obtained from the UCI machine learning repository, a self-organizing feature map (SOFM) and K-means algorithms were used. The authors achieved an accuracy of 97%.

This paper [3] proposed a classification framework that can be utilized in the firewall frameworks to create a legitimate classification for every transmitted packet by breaking down packet parameters by using shallow neural network (SNN) and optimizable decision tree (ODT) as machine-learning methods. In particular, the proposed models were utilized to prepare and group the Firewall dataset into three classes: "allow, "deny," and "drop/reset." The experiment scored an accuracy of 98%, and 99% for SNN, and ODT respectively.

[4] Analysed 500,000 instances using 6 features, which have been generated from Snort and TWIDS. The Action parameter was selected as the class attribute. The "Drop" and "Allow" parameters have been specified for the Action class. The firewall logs dataset was analysed and the features were inserted into machine-learning classifiers including Naive Bayes, kNN (k-Nearest Neighbours), One R and J48 using Spark in the Weka tool. The authors also compared the classification accuracy of these algorithms in terms of measurement metrics including Accuracy, F-measure and ROC values.

This experiment [5] proposed predictive models for predicting the work status at the finishing stage in the HPC framework. The model can be utilized as a device for checking the jobs in the HPC framework. The authors developed and built three models including HPC-CNN, HPC-AlexNet, and HPC-VGG16 based on the two machine learning techniques, which involved Initial and Transfer Learning of Convolutional Neural Network based on the HPC-work load dataset. Moreover, the three state-of-the-art Machine Learning methods: Artificial Neural Network (ANN), Classification and Regression Tree (CART) and Support Vector Machine (SVM) are used as the baseline models for performance comparison. The results show that the model that performs the best predictive performance is the proposed HPC-CNN model. The authors achieved 76.48% accuracy with the HPC-CNN

model followed with the CART model (75.60%), while the SVM model performs lowest the accuracy at 66.80%.

This study [6] classified some data from the Firewall Device used at a university, using a machine learning algorithm called multi-class support vector machine (SVM) classifier. As activation function for SVM classification they used linear, polynomial, sigmoid and Radial Basis Function (RBF). The authors measured the performance of the classifier by observing the estimation esteems of F1 score, recall and precision. The Action column is selected as the class attribute. The "deny", "drop", "allow" and "reset-both" attributes were implemented for the Action class. To obtain the maximum precision value in the classifier, SVM responses have been evaluated. The authors attempted to acquire the best activation function for the F1 score value. For each class, Receiver Operating Characteristic curves were also done.

III. METHODOLOGY

In this study, the Internet Firewall Dataset from UCI Machine Learning Repository was used [8]. The classification process is commonly done by coordinating the network packets against a set of guidelines and rules to block digital dangers from accessing the network. Subsequently, the firewall framework continues with either to "allow," "deny," or "drop/reset-both" the approaching packet. [1]

To classify the firewall log information, 11 of the characteristics in the informational index were chosen. While choosing information, it is critical to choose credits that have more numerical qualities. The action parameter has been acknowledged as a class. Table 1 shows the parameters and their description. [6]

Table 1. Dataset Columns and Description

Columns	Description
Source Port	The client source port number
Destination Port	The client destination port number
NAT Source Port	Network Address Translation Source Port Number
NAT Destination Port	Network Address Translation Destination Port Number
Action	allow, deny, drop, reset-both
Bytes	Total Bytes
Bytes Sent	Bytes Sent
Bytes Received	Bytes Received
Packets	Total Packets
Elapsed Time (sec)	Elapsed Time
Pkts_sent	Packets Sent
Pkts_received	Packets Received

There are 4 parameters in the action attribute used as a class. Descriptions of these parameters are shown in Table 2. [6]

Table 2. Dataset Action and Description

Action	Description
Allow	Explicitly allows traffic that matches the rule to pass
Deny	The firewall sends an ICMP type 3 (destination unreachable) message response back
Drop	The packet will be drop
Reset-both	A TCP reset is sent to both the client-side and server-side devices

Imbalanced classification represents a test for predictive modelling as the vast majority of the machine learning utilized for classification was planned around the assumption of an equivalent number of attributes for each class. This outcome in models that have poor predictive execution, explicitly for the minority class. As shown in Fig. 1, the dataset is imbalanced.

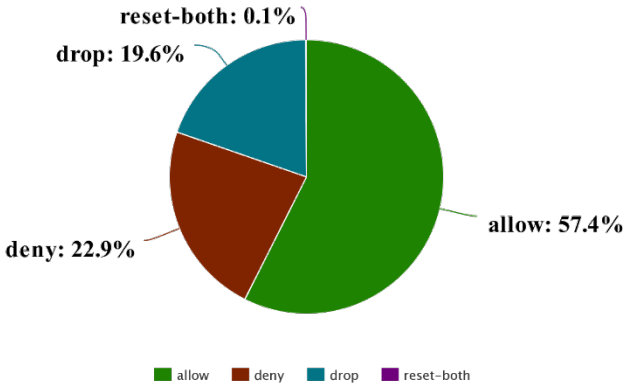

Fig. 1. Unbalanced Dataset

To counter this problem, the dataset needs to be balanced with resampling methods. The "reset-both" class has too few data to be re-sampled, so we merged "reset-both" and "deny" classes into one class. Undersampling was used to balance the uneven dataset by keeping each of the data in the minority class and diminishing the size of the larger part class.

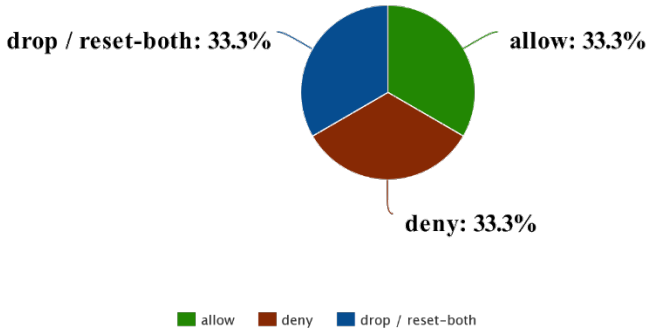

Fig. 2. Balanced Dataset

To have this prediction take place, we had to remap the variables of the Action column to 0, 1, 2 because the model does not work well with string values.

Table 3. Action and Mapped Value

Action	Mapped Value
Allow	0
Deny	1
Drop / Reset-both	2

For this model, an artificial neural network classifier with a standard scaler was used. Two hidden layers were configured with 10 and 5 neurons respectively. The L2 penalty was also configured to reduce the risk of overfitting.

The study was carried out on a laptop with specifications shown in Table 4. Anaconda [9] was used as software and scikit-learn [10] and imbalanced-learn [11] as main libraries.

Table 4. Laptop Specifications

CPU	Intel(R) Core (TM) i5-1035G1 CPU @ 1.00GHz, 4 Cores, 8 Logical Processors
GPU	Intel (R) UHD Graphics
RAM	8 GB DDR4
Hard Disk	1TB 5.4K RPM SATA Hard Drive

To build a more generalized model which can perform well on unseen data, k-fold cross-validation was used to partition the data. In this model, the data is split into 5 folds as this value have been shown exactly to yield test error rate estimates that experience neither from excessively high bias nor from very high variance

To avoid overfitting, a simple hidden layer was used with the default iteration and the L2 penalty parameter. The most common term for L2 penalty is L2 regularization. L2 regularization attempts to decrease the chance of overfitting by keeping the values of the weights and biases small.

To assess the performance of the models, four evaluators including accuracy, precision, recall, and F1-score were selected. All evaluators are figured from the confusion matrix table. [1]

The true positive (TP) is the number of the predicted data is “True”, and the actual data is “True”. The false-negative (FN) is the number of the predicted data is “False”, while the actual data is “True”. The false positive (FP) is the number of the predicted data is “True”, while the actual data is “False”. The true negative (TN) is the number of the predicted data is “False”, and the actual data is “False”. [1] [12]

The accuracy (1) is an evaluator that assesses the overall performance of the model. The recall (2) regards the model performance based on the actual value point view. Meanwhile, precision (3) observes the model performance based on the predicted value point of view. The F-measure or F1 score (4) is a harmonic mean of precision and recall. [1]

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{FN} + \text{TP} + \text{FP}} \quad (1)$$

$$\text{Recall} = \frac{\text{TP}}{\text{FN} + \text{TP}} \quad (2)$$

$$\text{Precision} = \frac{\text{TP}}{\text{FP} + \text{TP}} \quad (3)$$

$$\text{F}_1 = 2\,\text{X}\,\frac{\text{Recall X Precision}}{\text{Recall} + \text{Precision}} \quad (4)$$

IV. RESULTS AND DISCUSSION

Fig. 3 shows the confusion matrix which predicts an evaluation metric for the classification model of train values.

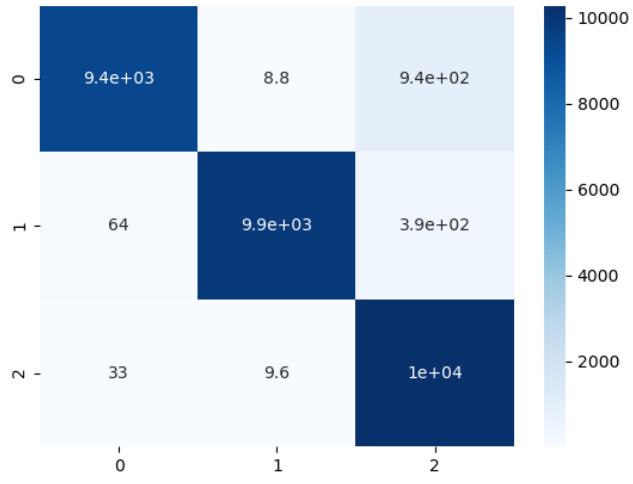

Fig. 3. Confusion Matrix for Train data

Table 5. Train Accuracy

Precision	Recall	F1-score	Accuracy
96.17%	95.81%	95.84%	95.81%

Fig. 4 shows the confusion matrix which predicts an evaluation metric for the classification model of test values.

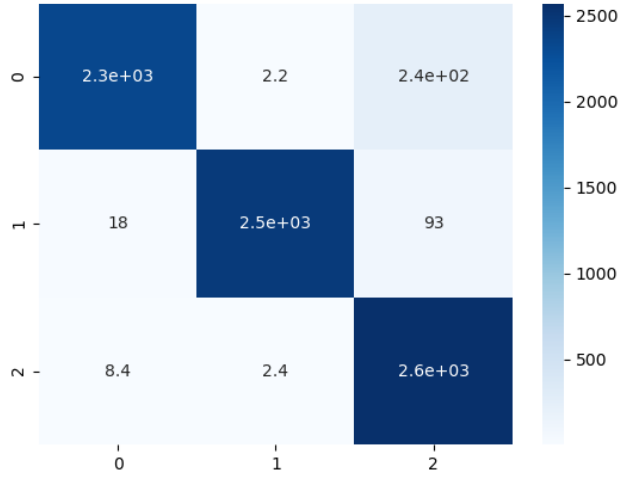

Fig. 4. Confusion Matrix for Test data

Table 6. Test Accuracy

Precision	Recall	F1-score	Accuracy
95.45%	94.93%	94.97%	94.49%

Fig. 5 shows the train v/s test accuracy during the 5-flod cross-validation.

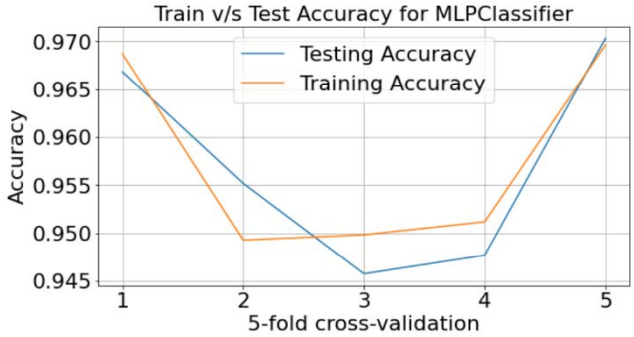

Fig. 5. Train v/s Test accuracy

Table 6 illustrates the test result for classifying the 3 classes log event. The model is giving accuracy up to 94.49%. Hence using our model, the network log data auditing and analysis can be done most optimally.

To acquire an understanding of the proposed model's benefits, our model was analysed by contrasting its accuracy with other best AI-based firewall-activity classification models in terms of classification performance metrics. The examinations are given in Table 7 beneath.

Table 7. Existing models and their accuracy

Research	Year	ML Technique	Accuracy
Fatih Ertam et.al [6]	2019	Support Vector Machine (SVM)	79.40%
Anupong Banjongkan et.al [5]	2020	Convolutional Neural Network (CNN)	76.50%
Shridhar Allagi et.al [2]	2020	Self-Organizing Feature Map (SOFM)	97.20%
Adrian Pıru et.al [7]	2019	Deep Neural Network (DNN)	92.82%
Our Model	2021	Deep Neural Network (DNN)	**94.49%**

Our model has a higher classification accuracy than most of the other existing related machine learning-based models.

V. CONCLUSION

The main objective of this paper was to develop a machine learning model, using Deep Neural Network representations. 11 features from the Internet Firewall Dataset from UCI Machine Learning Repository were used. As the dataset was imbalanced, undersampling method was used to create a generalized model. Our experiments made use of the k-folds technique and regularization to avoid overfitting. The proposed model has achieved 94.49% accuracy with testing data and 95.81% with training data. When comparing the accuracy of our approach with other research works, we noticed that ours performs better.

As future work, the plan is to integrate this trained model into a live system to better assess its generalization performance.

REFERENCES

[1] Qasem Abu Al-Haija and Abdelraouf Ishtaiwi,"Machine Learning Based Model to Identify Firewall Decisions to Improve Cyber-Defense," International Journal on Advanced Science, Engineering and Information Technology, vol. 11, no. 4, pp. 1688-1695, 2021. [Online]. Available: http://dx.doi.org/10.18517/ijaseit.11.4.14608.

[2] Shridhar Allagi, Rashmi Rachh. "Analysis of Network log data using Machine Learning" , 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), 2019

[3] Abu Al-Haija, Qasem & Ishtaiwi, Abdelraouf. (2021). Machine Learning Based Model to Identify Firewall Decisions to Improve Cyber-Defense. International Journal on Advanced Science, Engineering and Information Technology. 11. 1688. 10.18517/ijaseit.11.4.14608.

[4] As-Suhbani, Hajar & Khamitkar, S. (2019). Classification of Firewall Logs Using Supervised Machine Learning Algorithms. International Journal of Computer Sciences and Engineering. 7. 301-304. 10.26438/ijcse/v7i8.301304.

[5] A. Banjongkan, et. al., "A Comparative Study of Learning Techniques with Convolutional Neural Network Based on HPC-Workload Dataset" Inter. Journal of Machine Learning and Computing, vol. 10, no.1, 2020.

[6] Fatih Ertam, Mustafa Kaya. "Classification of firewall log files with multiclass support vector machine" , 2018 6th International Symposium on Digital Forensic and Security (ISDFS), 2018

[7] A. I. Pîrîu, M. Leonte, N. Postolachi and D. T. Gavrilut, "Optimizing Cleanset Growth by Using Multi-Class Neural Networks," in Proc. Of 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), Romania, pp. 425-429, 2018.

[8] Archive.ics.uci.edu. 2021. UCI Machine Learning Repository: Internet Firewall Data Data Set. [online] Available at: https://archive.ics.uci.edu/ml/datasets/Internet+Firewall+Data [Accessed 13 November 2021].

[9] Anaconda. 2021. Anaconda | The World's Most Popular Data Science Platform. [online] Available at: https://www.anaconda.com/ [Accessed 13 November 2021].

[10] Scikit-learn.org. 2021. scikit-learn: machine learning in Python — scikit-learn 1.0.1 documentation. [online] Available at: https://scikit-learn.org/stable/ [Accessed 13 November 2021].

[11] Imbalanced-learn.org. 2021. imbalanced-learn documentation — Version 0.8.1. [online] Available at: https://imbalanced-learn.org/stable/ [Accessed 13 November 2021].

[12] Kenta Sakamoto, Masaaki Okabe, Hiroshi Yadoshisa. "Generalized canonical correlation analysis for labeled data" , Procedia Computer Science, 2021

[13] Ranganathan, G., Fernando, X., Shi, F. and El Allioui, Y., 2021. Soft computing for security applications.