Dimensionality Reduction for Deep Learning Based Intrusion Detection Systems for IoT

Abayomi Ajiboye, Mosud Olumoye, Deborah Aleburu, Abisola Olayiwola, Dare Olayiwola, Samson Ajose

Abstract - Recent attacks on Internet of Things (IoT) devices have made it very important to provide a strong protection mechanism for IoT devices. A method of such protection mechanism is using Deep Learning models for Intrusion Detection Systems. Intrusion Detection Systems (IDS) can determine whether an invasion is currently ongoing or an intrusion has already occurred. In this study, the performance of the Deep Neural Network and Autoencoders were evaluated using the BoT-IoT dataset towards the development of binary classification model. The dataset was split into two; one for building the model and the other for validating it. The dataset was split into Normal- DoS, Normal-DDOS and Normal-Reconnaissance. This study has revealed that dimensionality reduction provided better classification of the dataset The model was developed for intrusion detection using DNN algorithms based on features extracted with and without Autoencoder for dimensionality reduction. The dimensionality reduction was performed to determine if it's possible to identify features that would improve the information required to effectively classify network intrusions. It was revealed that AeNN increased the accuracy of the classification of the Normal-DoS dataset for the classification of network intrusion detection by +3.4%, +11.5%, +13.1% for three simulations and a decline of -2.1% in the fourth simulation.

***Index Terms* - Auto-encoder, Dataset, Internet of Things, Intrusion Detection System, Deep Neural Network.**

I. INTRODUCTION

It is projected that the number and diversity of connected devices will grow exponentially with the development of the Internet of Things (IoT). While this offer users a variety of new applications and substantial benefits, it also opens up a number of new privacy, security, and safety threats, including risks to one's physical safety and personal security [1].

Intrusion Detection System (IDS), an important research accomplishment in the field of information security, can determine an invasion, which could be classified as an invasion currently going on or an intrusion that has already occurred. As a matter of fact, intrusion detection is tantamount to a classification problem be it binary or multiclass. Deep learning-based models for intrusion detection systems are used to improve the accuracy of classifiers in adequately identifying the intrusive behaviour [2].

According to [3], Intrusion Detection Systems leveraging the deep learning models are divided into three different classes as follows, generative, discriminative and hybrid. The generative models are those that use deep learning models for extracting features only and use shallow methods for classification. The discriminative learning models are those IDSs that use a single deep learning technique for both extraction and classification and the hybrid model is the IDSs that uses more than one deep learning method for generative and discriminative purposes. The generative model includes Deep Neural Network, Self-Taught Learning, Stacked Denoising Auto-Encoder etc. The discriminative models include Recurrent Neural Network, Convolution Neural Network etc.

The goal of this research is to evaluate the impact of dimensionality reduction on different deep learning models that would be better suited for Intrusion Detection Systems thereby providing far-reaching security for Internet of Things.

II. RELATED WORK

[4] posited that when a new technology is accepted by many people, there tends to be interest from cyber attackers who use different techniques in hacking. Because of the diverse types of IoT devices, the task of protecting these devices with a traditional IDS becomes a very tedious one. It was against this backdrop that [5] developed a novel ensemble of hybrid IDS for detecting attacks in IoT. This was done by fusing one class Support Vector Machine classifier and a C5 classifier and it was evaluated using the BoT-IoT dataset. They showed that combining two stages of the framework improves the detection accuracy. The result they obtained shows that the hybrid IDS performs better in terms of accuracy and false alarm rate relative to other machine learning techniques. [6] in their work, "Towards Deep learning Driven Intrusion Detection for the Internet of Things" developed a system of deep learning algorithms that can identify malicious traffic in IoT networks. They implemented the Deep Neural Network using Keras and

Manuscript received March 31, 2023; revised April 12, 2023.

A. Ajiboye is a M.Sc. student at Edinburgh Napier University, Merchiston campus, Scotland. ajiboyeabayomi@yahoo.com. Phone: +44 7767512597

M. Olumoye is a lecturer in the Department of Mathematics, Physics & Computer Science, Caleb University, Imota, Lagos, Nigeria. (myolumoye@yahoo.com)

D. Aleburu is a lecturer in the Department of Computer Science and Mathematics, Mountain Top University, Ibafo, Ogun State, Nigeria. (debaleburu56@gmail.com).

A. Olayiwola is a lecturer in the Department of Computer Engineering, Olabisi Onabanjo University, Ago-Iwoye, Ogun State, Nigeria.

D. Olayiwola is a lecturer in the Department of Computer Engineering, Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria.

S. Ajose is a postgraduate student of University of Lagos, Akoka, Lagos, Nigeria.

tested using the Cooja network simulator. The results were evaluated using the Texas Instruments Sensor tags CC2650. It was concluded from their work that after evaluating the performance of the model, precision rate of 95% and recall rate of 97% was recorded. Therefore, it is viable to use Deep Learning algorithms for intrusion detection in IoT.

[7] used a deep auto-encoder for intrusion detection system. The model was trained to avoid overfitting. The model created identifies normal and abnormal behaviour. The KDD-Cup ‘99 dataset was used in evaluating the model. The deep autoencoders were made up of four auto-encoders and each were trained using a greedy unsupervised layer-wise approach. The detection accuracy of the model was 94.71%.

[8] proposed a framework, Particle Deep Framework (PDF), a combination of Particle Swarm Optimisation (PSO) and deep learning. The framework was trained and validated using the BoT-IoT dataset. The PSO was used to select the hyperparameters of the Deep Neural Network. The PDF achieved a detection accuracy of 99.9% with false positives and negatives approaching zero. The performance of the deep learning models (Multi-layer Perceptron, Long Short-Term Memory, Convolutionary Neural Network and hybridized Convolutionary Neural network with Long Short-Term Memory) were compared and evaluated using the CICIDS 2017 dataset. The last layer of all the models was dense with sigmoid activation function. The maximum number of epochs was 100 because there was no improvement in the model beyond this. The hybridized model of CNN + LSTM had the highest accuracy of 97.16% while the MLP had the lowest, 86.34%.

[9] proposed a multi-layered recurrent neural network for implementation in fog computing security close to end users and IoT devices. The model was validated on NSL-KDD dataset. Mathew correlation and Cohen Kapp’s coefficients were added to the performance metrics of the model. The training algorithm was disintegrated into feed-forward computation, back propagation to output layer, back propagation to the hidden layer and weights update. The result revealed that the model showed high sensitivity to DoS attacks.

III. METHODOLOGY

Figure 1 shows the conceptual diagram of the research framework that was adopted in this study. The IoT dataset was obtained from the Research Cyber Range Lab of the UNSW Canberra which is made available at an online repository (https://cloudstor.aarnet.edu.au/plus/s/umT99TnxvbpkkoE). The dataset was initially subjected to data pre-processing which involved the purposive elimination of some features from the collected dataset. Following the process of the pre-processing of the IoT dataset collected, dimension reduction measures were adopted on the dataset. This was done to see if data generated from already-existing feature information may boost the performance of network intrusion detection classification based on the use of deep learning algorithms. The deep learning algorithms adopted were Deep Neural Network (DNN) which was adopted for supervised learning, and Autoencoder Neural Network (AeNN) which was adopted for dimensionality reduction and unsupervised learning of the BoT-IoT dataset.

Following this, the dataset was split into training and testing set based on a percentage proportion. A larger proportion of the IoT dataset was used for building the classification model based on the deep learning algorithms, while the smaller proportion was used to validate the classification model.

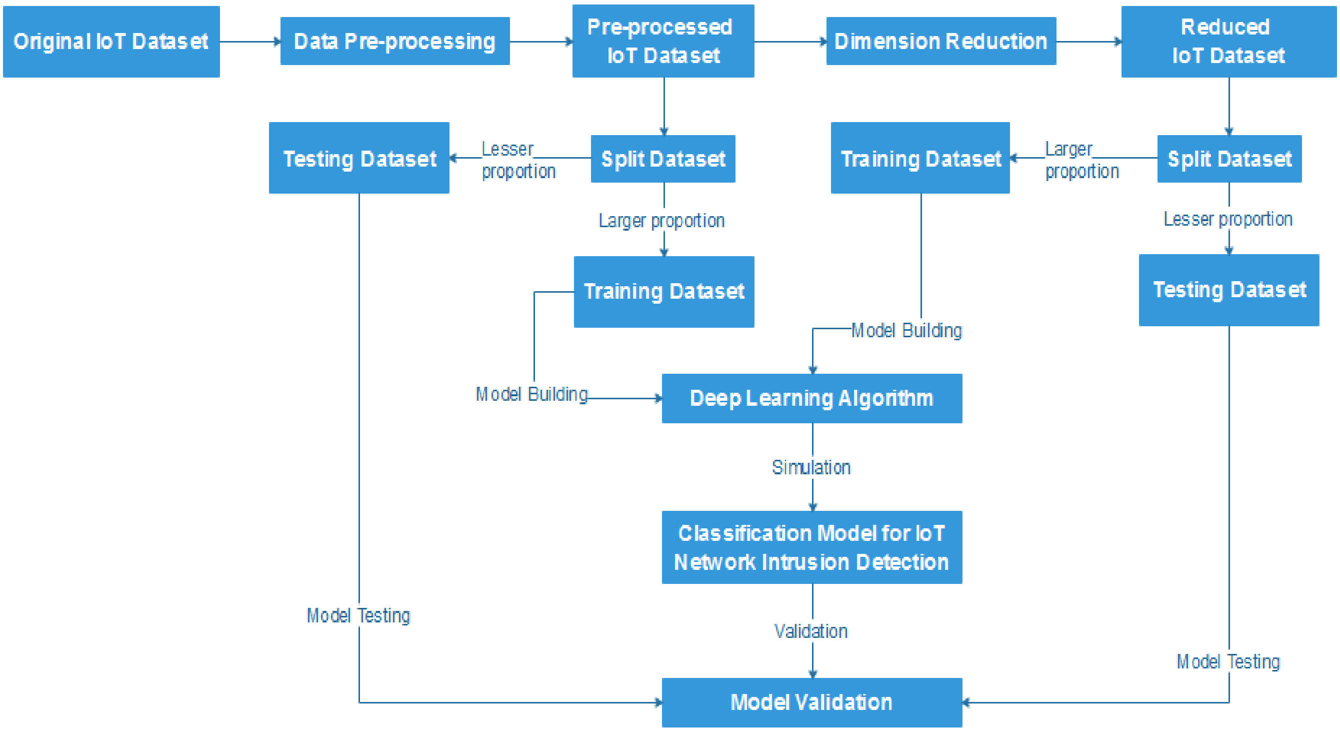

Fig 1: Conceptual Diagram of Research Framework

Several performance evaluation measures based on accurate and incorrect classifications were used to compare the performance of the classification model constructed using various percentage proportions and deep learning algorithms. Based on the models' performance assessments, the classification model with the best results was chosen. DoS, DDoS, and reconnaissance attacks are among the different attack traffic types considered in this study.

A. Model Building and Validation Process

In order to develop the classification model required for the detection of network intrusion, based on information collected from the botnet IoT dataset, there was a need to provide the process through which the building and validation of the model was performed. The percentage split technique was adopted for the training (building) and the testing (validation) of the classification model. This was done by allocating a larger percentage of the dataset for training (building) the classification model while the smaller percentage of the dataset was adopted for testing (validating) the classification model. Both categories of dataset were evaluated using accuracy, true positive, false positive and precision as evaluation metrics. However, the classification model with the best performance was selected during comparison based on the evaluation of the test dataset.

In this study, four (4) simulations were performed on the datasets that were generated using the deep learning algorithms that were proposed. Table I shows a description of the number of data records that were randomly selected for training and testing the classification model based on the datasets with binary and multi-class target variables.

Table I: Description of the Number of Records selected for Training and Testing

Simulation	Training Records	Testing Records
Simulation I (60%/40%)	571	381
Simulation II (70%/30%)	666	286
Simulation III (80%/20%)	761	191
Simulation IV (90%/10%)	856	96

B. Evaluation of classification model using performance metrics

The performance metrics that were used to evaluate the deep learning algorithms are presented in the following paragraphs.

a. Accuracy

This is defined as the proportion of total records which were correctly classified by the deep learning algorithm. The accuracy of a deep learning algorithm is expressed as a percentage (%). The accuracy of the classification model of the binary classifier is shown in equation (1).

$$Accuracy = \frac{A+D}{A+B+C+D} \times 100\% \qquad (1)$$

b. True Positive (TP) rate/Sensitivity/Recall

This is defined as the proportion of actual records that are correctly classified by a deep learning algorithm. It is used to determine how well a deep learning algorithm can recognize a class (the ability to distinguish between one class from the other). Equations (2a) and (2b) shows the expression for the TP rate of each of the class belonging to the binary classifier.

$$TP\ rate_{Normal} = \frac{A}{A+B} \qquad (2a)$$

$$TP\ rate_{Attack} = \frac{D}{C+D} \qquad (2b)$$

c. False Positive (FP) rate/False alarm rate

This is defined as the proportion of actual records that are misclassified by a deep learning algorithm. It is used to determine the inability of a deep learning algorithm to distinguish between one class from the other. The FP rate of the classification model of the binary classifier is shown in equations (3a) and (3b).

$$FP\ rate_{Normal} = \frac{C}{C+D} \qquad (3a)$$

$$FP\ rate_{Attack} = \frac{B}{A+B} \qquad (3b)$$

d. Precision

This is defined as the proportion of predicted records that are correctly classified by a deep learning algorithm. It is used to determine how correct the classifications made by a deep learning algorithm are. The precision of the classification model of the binary classifier is shown in equations (4a) and (4b).

$$Precision_{Normal} = \frac{A}{A+C} \qquad (4a)$$

$$Precision_{Attack} = \frac{D}{B+D} \qquad (4b)$$

IV. RESULT AND DISCUSSION

This section presents the result of the simulation of the classification model which was developed for the detection of network intrusion using the DNN algorithm based on the features that were extracted with and without AeNN for dimensionality reduction. Table II shows the summary of the result of the number of actual and predicted normal and attack records with and without AeNN.

Table II: Results of the Number of Actual and Predicted Records using DNN

Dataset	Simulation	Training	Testing	Without AeNN				With AeNN			
				Actual		Predicted		Actual		Predicted	
				Normal	Attack	Normal	Attack	Normal	Attack	Normal	Attack
Normal-DoS	Simulation I	571	381	190	191	207	174	185	196	187	194
	Simulation II	666	286	140	146	176	110	146	140	143	143
	Simulation III	761	191	95	96	122	69	102	89	100	91
	Simulation IV	856	96	51	45	52	44	40	56	41	55
Normal-DDoS	Simulation I	571	381	190	191	190	191	186	195	185	196
	Simulation II	666	286	140	146	140	146	132	154	132	154
	Simulation III	761	191	95	96	95	96	91	100	91	100
	Simulation IV	856	96	51	45	51	45	52	44	52	44
Normal-Reconnissance	Simulation I	571	381	190	191	229	152	186	195	186	195
	Simulation II	666	286	140	146	142	144	151	135	153	133
	Simulation III	761	191	95	96	94	97	93	98	93	98
	Simulation IV	856	96	51	45	51	45	51	45	53	43

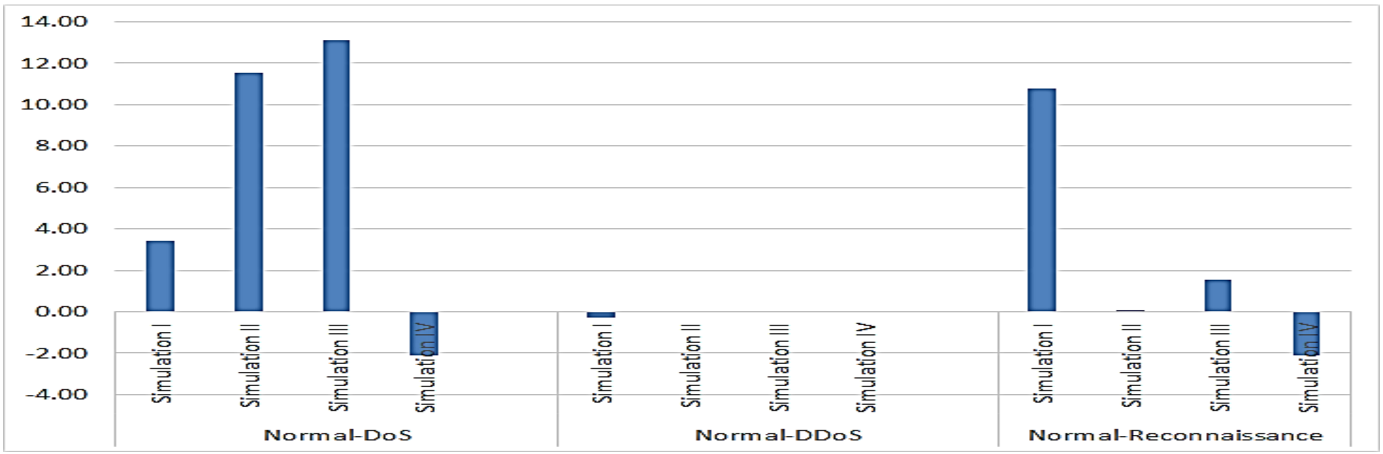

Fig. 2: Graphical Plot of the Change in Accuracy of Classification Model as a result of Dimension Reduction using AeNN

Figure 2 shows a graphical plot of the increase (bars above x-axis) and decrease (bars below x-axis) in the accuracies of the classification models developed using all the 3 datasets based upon the use of 4 simulations upon adopting the features selected by the dimension reduction performed by AeNN algorithm.

Figure 2 shows that the application of AeNN increased the accuracy of the classification of the Normal-DoS dataset based on Simulations I (+3.4%), II (+11.5%) and III (+13.1%) with the highest increase observed in Simulation III. However, there was a decline in the accuracy in Simulation IV (-2.1%).

The result in the diagram shows that the application of AeNN did not change the accuracy of the classification of the Normal-DDoS dataset based on Simulations II, III and IV which all had accuracy of 100% however there was a decline in the accuracy in Simulation I (-0.3%). The results in the diagram shows that the application of AeNN increased the accuracy of the classification of the Normal-Reconnaissance dataset based on Simulations I (+10.8%) and III (+1.6%) with the highest increase observed in Simulation I. However, there was a decline in the accuracy in Simulation IV (-2.1%) with no change in the accuracy of Simulation II.

The results showed that using the Normal-DoS dataset would require 90 percent of the dataset for training and 10 percent for testing without AeNN features, while using AeNN features would require 70 percent of the dataset for training and 30 percent for testing.

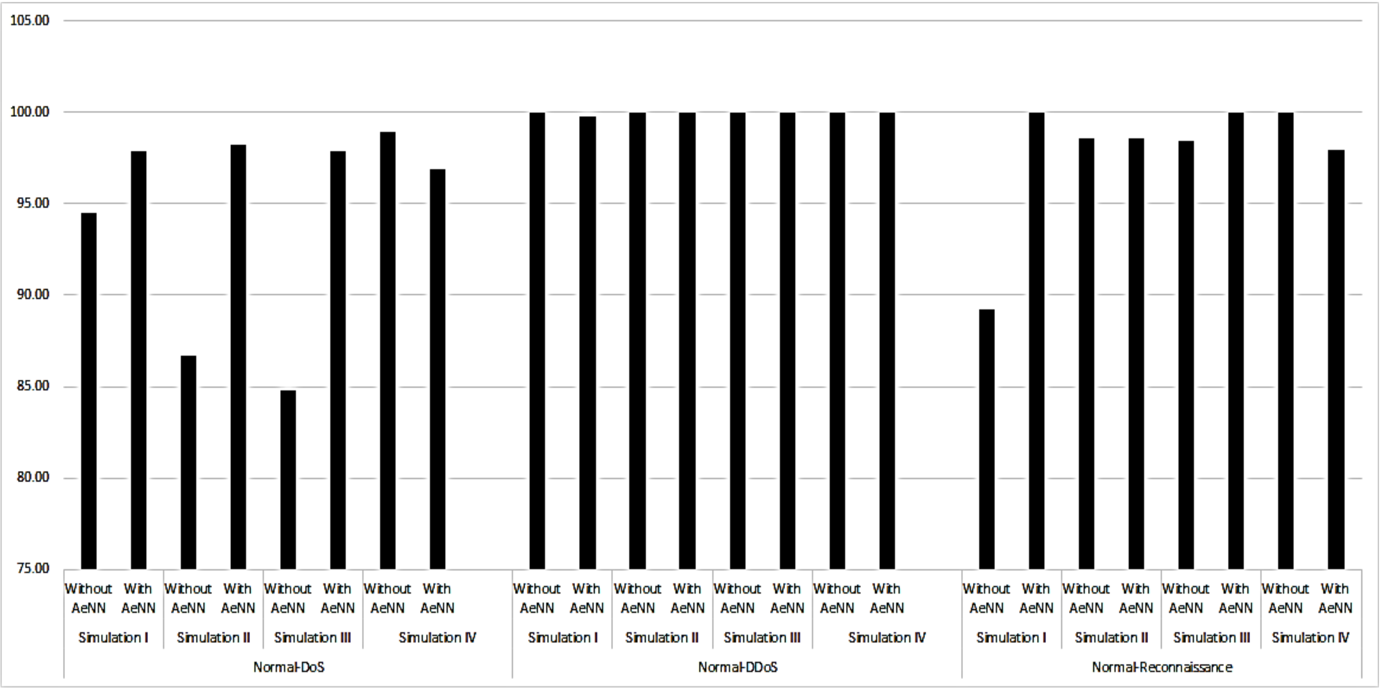

Fig. 3: Graphical Plot of Accuracies of BoT-IoT Dataset Simulations

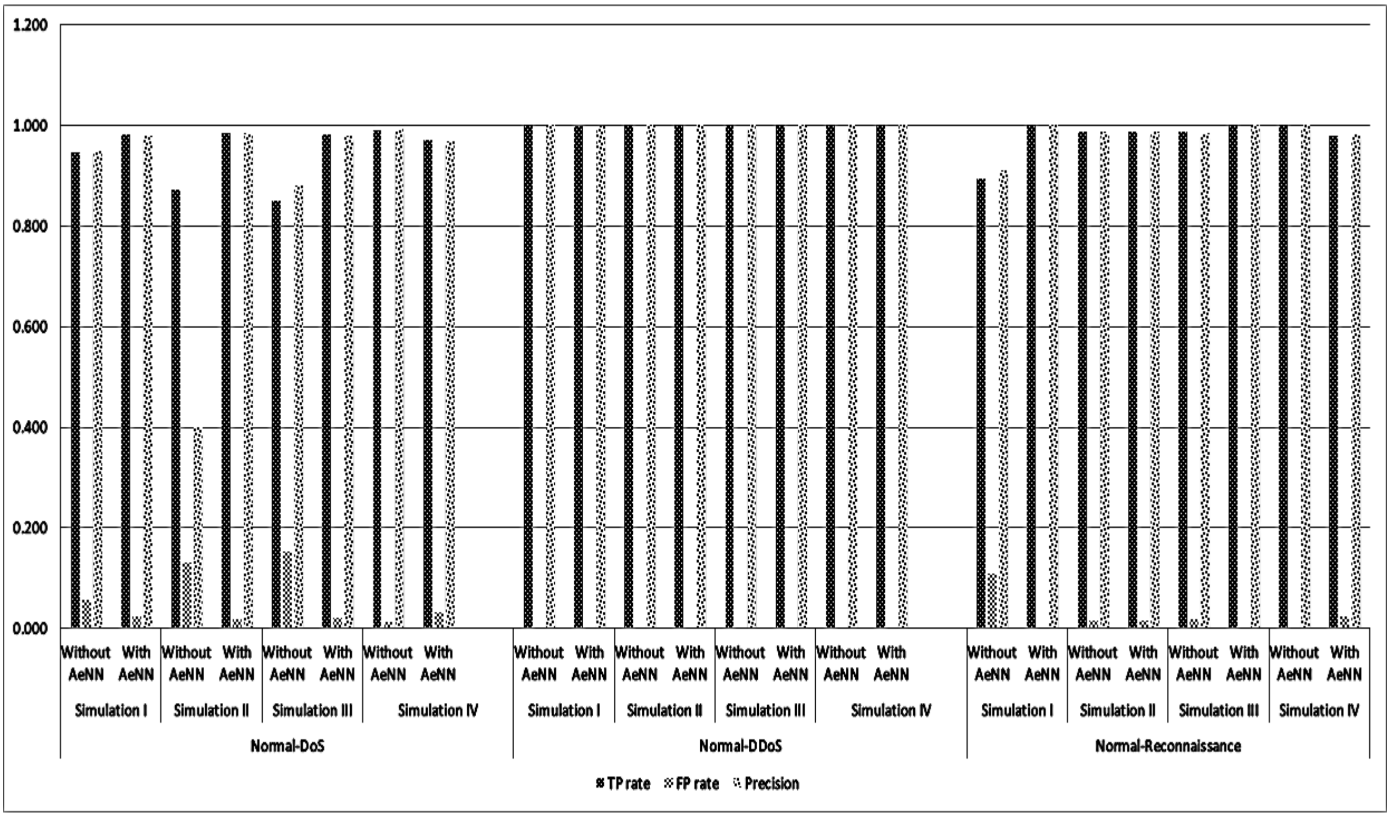

Fig. 4: Graphical Plot of TP rate, FP rate and Precision of BoT-IoT Dataset Simulations

Figure 3 shows the graphical plot of the accuracy of the 4 simulations performed on the 3 set of datasets selected for this study while Figure 4 shows the graphical plot of the TP rate, FP rate and precision of the classification models. The diagrams show the relative effect of the application of dimensionality reduction on the accuracy, TP rate, FP rate and precision of the classification models created for each class of BoT-IoT dataset selected for this study.

V. CONCLUSION

The study developed binary classification model for the detection of network intrusion using deep neural networks (DNN) using data containing features extracted with and without the use of Autoencoder neural network (AeNN). The study compared the performance of the classification models in order to determine the relative impact of dimensionality reduction using AeNN on the performance of

DNN for the binary classification network intrusion detection using information collected from BoT-IoT dataset.

REFERENCES

[1] O. Brun, Y. Yin, J. Augusto-Gonzalez, M. Ramos, and E. Gelenbe, "IoT Attack Detection with Deep Learning*". ISCIS Security workshop, 2018.*

[2] N. Chaabouni, M. Mosbah, A. Zemmari, C. Sauvignac and P. Faruki. "Network Intrusion Detection for IoT security based on learning techniques" in *IEEE Communications Surveys & Tutorials,* 2020.

[3] K. Kim, M. Aminanto. H. Tanuwidjaja, "Network Intrusion Detection using Deep Learning: A feature approach" in *Springers Briefs on Cyber Security Systems and Networks,* 1st ed., 2018.

[4] N. Koroniotis, N. Moustafa, E. Sitnikova and B. Turnbull, "Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: BoT-IoT dataset" in *Future Generation Computer Systems*, 2019.

[5] A. Khraisat, I. Gondal, P. Kamruzzaman and A. Alazab, "A novel ensemble of hybrid intrusion detection system for detecting internet of things attack" in *Electronics (Switzerland),* 2019.

[6] G. Thamilarasu and S. Chawla, "Towards deep-learning -driven intrusion detection for the internet of things" in *Sensor (Switzerland)* 2019.

[7] F. Farahnakian and J. Heikkonen, "A Deep Auto-encoder based approach for intrusion detection system" in *International Conference on Advanced Communication Technology,* 2018.

[8] N. Koroniotis and N. Moustafa, "Enhancing Network Forensics with Particle Swarm and Deep Learning: The Particle Deep Framework" in *7th International Conference on Artificial Intelligence and Applications,* 2020.

[9] M. Almiani, A. AbuGhazleh, A. Al-Rahayfeh, S. Atiewi and A. Razaque, "Deep Recurrent Neural Network for IoT Intrusion Detection System" in *Simulation Modelling Practice and Theory,* (2019).

[10] M. Roopak, G. Tian and J. Chambers, "Deep Learning Models for Cyber Security in IoT Networks" in 2019 IEEE 9th Annual Computing and Communication Workshop and Conference.