Science World Journal Vol. 17(No 1) 2022
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

UNCERTAINTIES CLASSIFICATION IN CYBERSPACE USING ENSEMBLE LEARNING MODEL

*M.E. Irhebhude, Z.O. Musa and A.O. Kolawole

Department of Computer Science, Nigerian Defence Academy, Kaduna

*Corresponding Author Email Address: mirhebhude@nda.edu.ng

ABSTRACT:
Studies have shown that different techniques can help classify uncertainties in the cyber space, however, a lot of these studies did not report the false predictions. Ensembled classifier was applied in this paper to curb the uncertainties in cyberspace. Classification learner in MATLAB was used as a tool to train the machine learning model on the publicly available University of New South Wales Network-Based (UNSW-NB15) and locally gathered datasets. A multiclass classification was done on the two datasets which consist of various attack categories. An experiment was performed with the proposed model on the datasets with the use of an ensemble classifier in MATLAB classification learner with 30% held for validation. Performances were measured using accuracy, confusion matrix, and receiver operating characteristics (ROC) curve. The experiments resulted in excellent classification accuracy of 99.1% and 99.4% on the merged Comma Separated Value (CSV) UNSW-NB15 dataset and self-acquired dataset respectively. Experimental results from the two datasets have shown that ensembled gave a more robust classification accuracy compared to artificial neural network classifier. With the results, the ensemble model help solved the problem of classification of attacks in network environment and uncertainties in cyberspace. Infrastructures in cyberspace and user interaction will be well secured with the adopted solution.

Keywords: UNSW-NB15, Network Intrusion Detection Systems, Classification, Ensembled Classifier

INTRODUCTION

A rise in internet attacks has led to various security measures put in place to mitigate and minimize various attacks. Intrusion Detection Systems (IDS) used with firewalls act as supplement to monitor and analyze security threats and violations (Mazini et al., 2018).

A lot of research has been done on applying network anomaly detection in various environments ranging from aircraft engine measurements, cloud data center temperature, to telecommunication and Automated Teller Machine (ATM) fraud detection (Nawir et al., 2018). In the development of efficient IDS models, a large amount of data is required for training and testing. There are many publicly available datasets derived from the misuse-based and anomaly-based approach for research in network intrusion detection systems. These researches have been carried out using Machine Learning (ML) and Deep Learning (DL) techniques (Al-Daweri et al., 2020).

This paper focuses on generating an intrusion detection-based dataset from penetration testing sessions, classifying the dataset using an ensemble classification technique. The University of New South Wales Network-Based (UNSW-NB15) was used as benchmark dataset technique validation. The self-generation dataset was done to see if variant attributes from the benchmark dataset can effectively predict the selected attacks.

The UNSW-NB15 data is a publicly available dataset that was created using the Perfect Storm tool in the Cyber Range Lab of the Australian Center for Cyber Security (ACCS). The dataset is sub-divided into four Comma Separated Value (CSV) files and contains normal and modern-day attack traffics. It contains nine attack categories which include; Analysis, backdoor, Denial-of-Service (DoS), Exploit, Fuzzers, Generic, Reconnaissance, Shellcode, and Worms (Moustafa and Slay, 2016).

The penetration testing dataset was collated from a session of synthetic attacks on Metasploitable 2, using Kali Linux. Metasploitable 2 is an intentionally vulnerable Ubuntu Linux virtual machine that provides a secure environment for penetration testing. Kali Linux is a Debian-derived distribution designed for digital forensics and penetration testing. Wireshark is an open-source packet analyzer tool used for network analysis. All information about packet exchange between these Kali Linux and Metasploitable 2 were recorded using Wireshark.

Studies have shown that a lot of instances (high volume dataset) can help to adequately classify uncertainties in cyberspace (Nawir et al., 2018), when the volume of data is in millions, results also show that a lot of time is required for prediction to take place accurately. Reported works have failed to show false predictions despite the excellent accuracy score obtained (Srinivasan et al., 2020). Ensemble and Artificial Neural Network (ANN) classifiers have been used by various authors to predict uncertainties in cyberspace. This study uses an ensemble classifier to predict uncertainty in cyberspace and reports the impact it has on false prediction in comparison with the ANN classifier. It is reasonable to believe that having several classifiers 'working together can have the potential to give better predictive accuracy than one on its own (Bramer, 2013), hence the choice to use ensemble in this study.

The paper is organized as follows: Section 1 introduces the core concept while section II discusses the ensemble classifier. Section III is the review of related works. The proposed methodology and detailed experiment carried out are shown in section IV while section V contains analysis and discussion of results as well as a comparison with an existing model. Section VI concludes the study.

ENSEMBLE CLASSIFIER

Ensemble learning involves the training of collection of a different number of classifiers to perform classification tasks in other to achieve good performance beyond using individual classifiers. (Yang, 2011). It involves training several neural networks and then combining the components prediction in solving the problem, There are several approaches in neural network ensemble for training components such as Boosting and Bagging (Zhou et al., 2002).

Boosting proposed by Schapire (1990) and improved by (Freund, 1990) creates a series of component neural networks whose

training sets are defined by their performance. In the work of Bramer (2013)ensemble classification learns several classifiers, and then combines their predictions with the hope to give a more robust classification of unseen instances using a voting technique (See Figure 1). If the combined classifiers are of the same types the ensemble is known as homogeneous, otherwise heterogeneous.

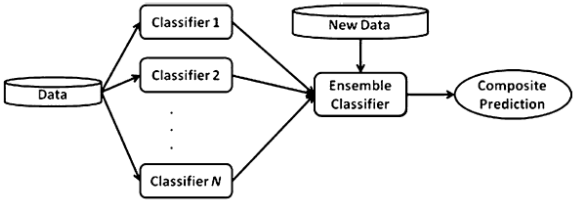

Figure 1: Ensemble Classification

Homogeneous ensemble formation using decision trees example can be seen as:

i. M trees are generated using the same tree generation algorithm, with different parameter settings, all using the same training data.
ii. M trees are generated using the same tree generation algorithm, all with different training data and either with the same or with different parameter settings.
iii. M trees are generated using a variety of different tree generation algorithms either with the same or with different training data.
iv. M trees were generated using a different subset of the attributes for each one.

An ensemble classification algorithm according to(Bramer, 2013) is:

1. Generate *K* classifiers for any dataset
2. For the new dataset *Y*
   a) Predict *Y* for each of the *K* classifiers
   b) Select the classifier with the most prediction strength. This is called a majority voting model where each time a classifier gives a correct prediction for an unseen instance it counts as one 'vote' for that classifier.

In this study, MATLAB was used for the experiment using the classification learner app. The ensemble classifier with the following search range and parameters were selected and used for the experiment.

a) Ensemble methods: Bag, AdaBoost, RUSBoost
b) Number of learners: 10-500
c) Learning rate: 0.001-1
d) Maximum number of splits: 1-1400001

For the validation, the following parameters were selected by the classifier and used for the classification.

a) Ensemble method: AdaBoost
b) Maximum number of splits: 205
c) Number of learners: 479
d) Learning rate: 0.92058

RELATED WORKS

Studies show that ANN and ensemble classifiers have been used as a classifier for the classification of uncertainty in cyberspace using the UNSW-NB15 dataset as a benchmark. The focus of the review in this study will be on studies that used ANN and ensemble as a classifier with the UNSW-NB15 dataset.

A novel approach involving a 6-step algorithm using chaos theory and ANN was proposed in (Aljumah and Ahamad, 2016). The 6-step algorithm includes; gathering network data, data preprocessing using aggregate averaging, prediction of traffic, determining prediction error, detecting attack traffic using chaos theory, and using ANN to detect DDoS. This model used a locally gathered dataset from network analysis which involved traffic from synthesized DDoS attacks. The dataset was huge and had to be subdued by averaging sequence with time session. The authors used statistical functions for predicting the Lyapunov constant to evaluate predicted error and differentiate between genuine and attack traffic. The unsupervised learning technique used was the clustering technique and it divided the data into 3 clusters, burst, genuine and DDoS traffic. These clusters were then used with supervised learning to reduce the error of backpropagation. This resulted in a 95% accuracy rate of detection of DDoS attacks but did not report result for false predictions

Idhammad et al. (2017) proposed a Feed-Forward Neural Network (FNN) and ANN-based Detection Method (ADDM) to detect DoS attacks. Experiment with the UNSW-NB15 and NSL-KDD datasets were used to test the model's performance. Consistency-based Feature Selection and Correlation-based Feature Selection (CFS) were used to filter the features after they were ranked statistically. The Pearson Correlation Coefficient (PCC) which is a measure of dependency between variables was used in the CFS approach. The normal and attack (DoS) traffic were separated from the datasets and labeled as 1 for DoS attack and 0 for normal traffic. The MLP algorithm was used for the classification of these datasets. Backpropagation algorithm was also used with Stochastic Gradient Descent (SGD). Multilayer Perceptron (MLP) and ADDM were trained and tested using the extracted dataset that contained only normal and DoS traffic. The performance metrics used were accuracy, sensitivity, specificity, false alarm rates, processing time, ROC curves, and AUC values. The performance of ADDM was compared to that of unoptimized MLP (u-MLP), NSL-ANN, HSV-ANN, DDMA, and ANN. ADDM resulted in an accuracy, testing time, sensitivity, specificity and FAR values of 97.1%, 0.46 seconds, 97%, 100% and 0.06% with UNSW-NB15 dataset and 99.2%, 0.35 seconds, 99%, 100% and 0.02% with NSL-KDD dataset respectively while u-MLP resulted in 79.2%, 3.05 seconds, 82%, 87% and 0.14% with UNSW-NB15 dataset and 83.5%, 2.16 seconds, 90%, 93% and 0.11% with NSL-KDD dataset respectively. The ADDM and u-MLP had the best performance compared to other models but recorded high false positive rates.

A binomial classifier for NIDS was proposed in (Al-Zewairi et al., 2017). Three experiments were carried out on the UNSW-NB15 dataset to ascertain the optimal activation function, select the primary features and test the proposed model on unseen data. The proposed model was built using a native implementation of multi-layer feed-forward ANN using backpropagation and SGD. The first experiment aimed to find out the optimized activation function, the second experiment involved training the model with the best activation function resulting from the first experiment to choose the primary features and the third experiment was executed using the result of the first two experiments. The model was tested on unseen data which was broken into three subsets; 60% training, 10% validation, and 30% testing. The performance metrics were accuracy, F1 score, FAR, specificity, Area Under Curve (AUC), precision, recall, and training time. In terms of accuracy and FAR values, the proposed model reported an outstanding result with the

highest accuracy value of 98.99%, and the lowest FAR value of 0.56 compared to the values of other techniques.

A comparison of the performance of ANN and NB algorithms on a dataset collected from Ahmad Dahlan University research laboratory using Wireshark was carried out in(Yudhana et al., 2018). The features were selected using statistical methods. The dataset was trained using MATLAB. The trained data was made up of 70 DDoS data and 30 normal traffic. Testing of the ANN classifier was done using 20-log data which gave 95.23% accuracy. The NB Gaussian method was used to test the 20-log data and it resulted in an accuracy rate of 99.9%. The performance of the classification algorithms was measured by the accuracy rate and NB was shown to have performed better than ANN. The study concluded that NB is better than ANN and recommended that sample size, variations of the hidden layer as well as other classifications such as SVM be considered in future works to get improved accuracy.

An investigation was carried out on the outcome of binary and multi-class classification on the UNSW-NB15 and CICIDS2017 datasets using RF, GBT, and Deep Feed-Forward Neural Network (DFNN) classifiers in(Faker and Dogdu, 2019). Homogeneity metric was used for feature selection and the experiment was carried out using 5FCV. The CICIDS2017 dataset was pre-processed and this resulted in the Rep-CICIDS2017 and Rem-CICIDS2017 dataset. Classifying the UNSW-NB15 dataset, RF had an accuracy of 98.85% and 98.86%, GBT had an accuracy of 97.83% and 97.92% while the DFNN had an accuracy of 99.16% and 99.19% using the complete dataset and selected features for binary classification respectively. With the complete dataset and selected features for multi-class classification, RF had an accuracy of 91.76% and 91.77%, and DFNN had an accuracy of 97.01% and 97.04%. Classifying the Rem-CICIDS2017 dataset, RF had an accuracy of 92.54% and 92.71%, GBT had an accuracy of 99.81% and 99.99% while DFNN had an accuracy of 97.71% and 97.72% using the complete dataset and selected features for binary classification. With the complete dataset and selected features for multi-class classification, RF had an accuracy of 92.54% and 92.71% and DFNN had an accuracy of 99.56% and 99.56%.

A binary classification was conducted in (Kanimozhi and Jacob, 2019) to experiment on the training and testing subsets of the UNSW-NB15 dataset. This research performed feature selection on the dataset using the Recursive Feature Elimination (RFE) approach with RF classifier and created an ANN with Anaconda, Jupyter Notebook, and Scikit-Learn using Python 3. The proposed technique resulted in an AUC of 0.98, and the classification of the testing and training sets were done independently. The classification of the training set resulted in a 96% accuracy, 97% precision, 96% recall, F1-score of 97%, and AUC of 99% while the classification of the testing set resulted in 89% accuracy, 99% precision, 85% recall, F1-score of 91%, and AUC of 98%. This study mapped outsource time to live (sttl), several connections with the same source and destination address in the last one hundred records according to the last time (ct_dst_src_ltm), source to destination bytes (sbytes), and source load (sload) as the top four features on the dataset as the classification with these features resulted in the highest accuracy of 98.3%.

A survey on ML techniques was conducted in(Seraphim et al., 2019). The paper applied some ML algorithms on the NSL-KDD dataset and compared these algorithms based on accuracy, f-measure, confusion matrix, and recall. The proposed model was based on ANN and the result was compared to that of DT, Simple Logistic Regression (SLR), and KNN classifiers. In data pre-processing, the last 9 features of the dataset were dropped and the categorical attributes were encoded using the LabelEncoder feature of the Scikit Learn package. The model was built based on KNN and SLR classifiers. The experiment was done using 10FCV. The training set was made up of 4 attack types, DoS, probe attack, User-to-Root (U2R), and Root-to-Local attacks (R2L). From the results of the experiment, ANN had the highest accuracy of 99.46% though the precision varied with tuning parameters.

The evaluation of a classification model on the UNSW-NB15 and NSL-KDD datasets was carried out in (Elkassabi et al., 2020) using the Waikato Environment for Knowledge Analysis (WEKA) as a simulation tool. Feature selection was done with the use of Correlation-based Feature Selection (CFS) and Information Gain (IG) for feature selection. A subset of both datasets was used for classification as WEKA cannot handle large datasets. 20% of the NSL-KDD dataset and ten thousand (10,000) instances of the UNSW-NB15 dataset were extracted to experiment. A binary classification was carried out on the UNSW-NB15 dataset and resulted in an F-measure of 98.8% with CFS, 91.3% with IG, and 99.8% using IG and CFS. Classifying the NSL-KDD dataset resulted in an F-measure of 90.8% with CFS, 91.8% with IG, and 95.2% using IG and CFS.

Authors in (Awujoola et al., 2021) proposed a combination of wrapper feature selection based on genetic algorithm with a combination of Synthetic Minority Oversampling (SMOTE) and resampling technique for intrusion detection, the KDDCUP99, and NSL-KDD datasets were used on three different decision three classifiers to achieve accuracy of 99.9873% and 99.8457% respectively.

An investigation on the performance of DL with an improved version of the UNSW-NB15 dataset within two classification categories, binary and multi-class classification was carried out by (Aleesa et al., 2021). This research compared the proposed model used to models in previous works to evaluate the efficiency of DL and ML models. This dataset was divided into70% training, 15% for testing, and 15% for validation. The binary and multi-class classification was done using ANN, DNN, and RNN with LSTM, RNN-LSTM. Using ANN, the hyperparameters specified were hidden layer, number of neurons, optimizer, hidden layer activation function, Output layer activation function, epochs, and batch size with values 1, 850, Adam, ReLU, SoftMax, 100, and 100 respectively. Binary classification yielded an accuracy of 99.26%, 1.51% loss while multi-class classification yielded an accuracy of 97.89%, and 5.27% loss. Using DNN, the hyperparameters specified were hidden layer, number of neurons, optimizer, hidden layer activation function, Output layer activation function, epochs, and batch size with values 3, 100, Adam, ReLU, SoftMax, 100, and 100 respectively. Binary classification yielded an accuracy of 99.22%, 1.56% loss while multi-class classification yielded an accuracy of 99.59%, and 0.92% loss. Using RNN-LSTM, the hyperparameters specified were hidden layer, number of neurons, optimizer, hidden layer activation function, Output layer activation function, epochs, and batch size with values 3, 128, 64, and 32 neurons at each layer, Adam, ReLU, SoftMax, 100 and 100 respectively. Binary classification yielded an accuracy of 85.42%, 35.18% loss while multi-class classification yielded an accuracy of 85.38%, and 48.56% loss. However, the paper did not report details of the confusion matrix or False Positives (FP).

Studies in (Rashid et al., 2022) and (Sharma and Yadav, 2021) adopted the ensembled classifier for the prediction of uncertainties using the UNSW-NB15 dataset. However, the authors classified

the dataset using binary classification. Similarly, the authors did not clearly show false predictions obtained from the study.

Authors in (Aleesa et al., 2021) recommended that further findings be done by deploying the framework in a real environment. Also, from reviewed works of literature, the majority of studies did not report FPR and FNR performance values. The impact of iteration limits was also not investigated in most studies. Studies did not clearly show works that adapted ensembled classifiers for the classifier of attack categories of the UNSW-NB15 dataset. Therefore, this study will classify uncertainties in cyberspace using an ensembled classifier on two datasets; UNSW-NB15 and locally generated data using slightly different and fewer attributes between datasets. The UNSW-NB15 will serve as a benchmark and all categories of attacks will be considered for the experiment.

MATERIALS AND METHOD

The proposed methodology has various stages which includes dataset collection, data preparation, and classification, partitioning of data, and result.

Datasets

The experiment used two datasets that were available online (UNSW-NB15) and locally sourced.

The UNSW-NB15 Dataset

The UNSW-NB15 is a publicly available online dataset created by the University of New South Wales, UNSW in 2015 using the Perfect Storm tool in the Cyber Range lab of the Australian Center for Cyber Security, ACCS (Moustafa and Slay, 2016). The datasets are in four CSV files, these files were combined in a single CSV file totaling 2, 540,044. However, due to the limit of the MATLAB environment, the entire sets were not considered for the experiment. Similarly, for clarity of approach, the first two CSV files were combined to have 1,400,002 instances. This dataset has nine families of attacks, namely, Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, and Worms. The Argus, Bro-IDS tools were utilized and twelve algorithms were developed to generate 49 features used as class labels.

Local Dataset

The penetration testing dataset used was obtained from a cumulative report of observations from network analysis carried out during penetration testing. The network analysis was carried out using Wireshark, an open-source packet analyzer while Kali Linux was used to attack the Metasploitable 2. Kali Linux is a Debian-derived Linux distribution designed for digital forensics and penetration testing while Metasploitable 2 is an intentionally vulnerable machine Ubuntu Linux virtual machine that provides a secure environment for penetration testing and security research. The dataset has twenty-nine (29) features and comprises normal traffic and three (3) different types of attacks namely; reconnaissance, exploit, and backdoor attacks. Penetration testing was the first step of the experiment carried out using Kali Linux and Metasploitable 2 as two virtual machines set up on VMware. Kali Linux and Metasploitable 2 were configured to have a bridged network and the communicating interfaces, Ethernet port zero, (Eth0) on both machines were set to have IP addresses of the same network. Eth0 on Kali Linux had an IP address of 192.168.56.3/24 while Eth0 on Metasploitable 2 had an IP address of 192.168.56.4/24. Kali Linux was used to carry out the penetration testing on Metasploitable 2 using two types of attacks, Very Secure File Transfer Protocol Daemon (VSFTPD) version 2.3.4 and Samba attack.

The exploit attack was generated using the Samba exploit in Kali Linux while the backdoor attack is the Very Secure File Transfer Protocol Daemon (VSFTPD) backdoor. The vulnerability assessment conducted for the VSFTPD attack was a port scan using the Network mapper (Nmap) command on the Kali Linux terminal. The Nmap command scanned through the ports on the target machine, Metasploitable 2, and displayed a list of port numbers, states, and services. The state specified whether the port is open or not while the service showed the protocols on the ports. On the completion of this task, FTP port 21 was found open and the attack was launched specifying the IP address of Metasploitable 2 as the IP address of the target machine. There was a remote login from the host, Kali Linux to the target machine as a "root user", a user that has access to the file directory of the machine. A new directory was created on the target machine, a file was created in that directory and saved on the target machine.

Samba attack was also launched using Kali Linux as the host machine and Metasploitable 2 as the target machine. Samba is an implementation of the Server Message Block (SMB) protocol that is implemented on Windows and Linux systems.

Wireshark was used to capture the packet exchange between the host and target machines while VSFTPD and Samba attack sessions were in progress. The Wireshark recordings for each session were saved in packet capture (PCAP) file on the virtual machine. The PCAP files gotten from Wireshark were converted to CSV and the files were combined to form a single CSV file totaling 5,534.

Data Preparation and Classification

Data preparation was done on Microsoft Excel as the CSV files containing the datasets had columns required to be filled manually. As a result, the datasets in the CSV files were ready to be fed into the ensembled model. Figure 2 shows a block diagram of the proposed model; import dataset, classification learner, partition, and result. A new MATLAB workspace was opened and the CSV reader was fed into the classification learner to train the model with 30% held for validation. The output of the experiments was displayed with accuracy, confusion matrix, and Receiver Operating Characteristics Curve (ROC).

Science World Journal Vol. 17(No 1) 2022
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

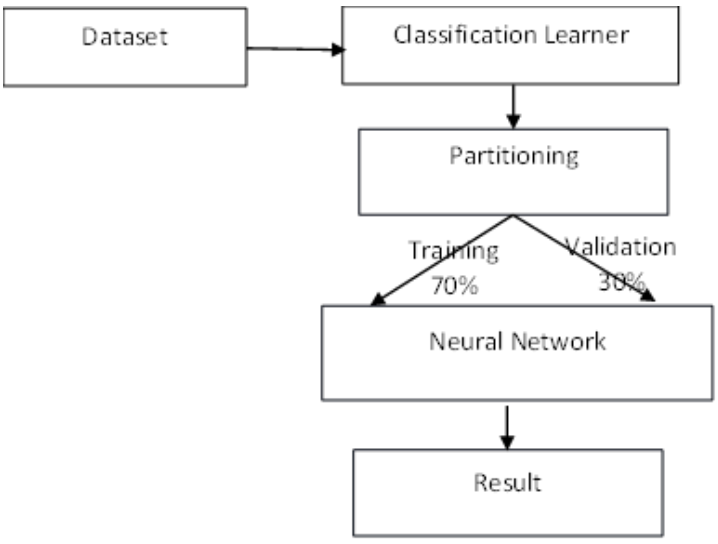

Figure 2: Flow diagram of Proposed Model

RESULTS AND DISCUSSION

Experiments were conducted on the UNSW-NB15 dataset and local dataset, the performance of the proposed model was evaluated by several criteria which include True Positive Rate (TPR), False Negative Rate (FNR), Positive Predictive Value (PPV), Confusion Matrix, and Receiver Operating Characteristics (ROC) Curve.

Analysis of the UNSW-NB15 dataset

Executing the classification model with the UNSW-NB15 dataset resulted in an accuracy of 99.1% with a training time of 56235secs. The Ensemble classifier was used for classification with the following model type settings; Preset: Optimizable Ensemble, Learner Type; Decision tree. The result of the experiments conducted is displayed in the confusion matrix and ROC curve as shown in figures 3 to 6.

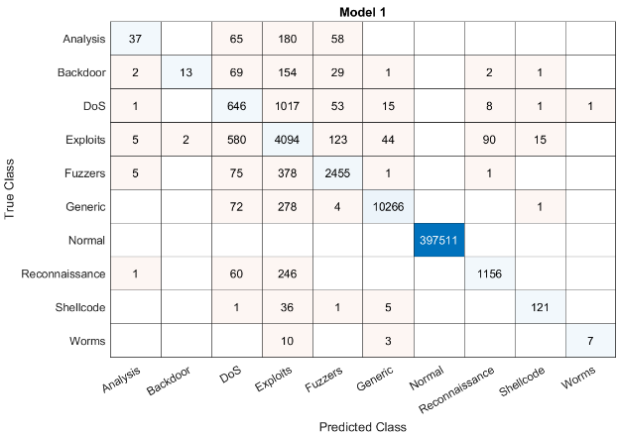

Figure 3: Confusion Matrix showing Number of Observations on UNSW-NB15 Dataset

Figure 3 shows the number of observations obtained in classifying the various categories of attack. The classifications were in ten categories which include: Analysis, Backdoor, DOS, Exploits, Fuzzers, Generic, Normal, Reconnaissance, Shellcode, and Worms all having correctly predicted values of; 37, 13, 646, 4094, 2455, 10266, 397511, 1156, 121 and 7. A total of 1929 and 1765 FPR and FNR were recorded. The high FPR indicates a higher rate of the categories of attacks to be wrongly classified to belong to other classes.

The confusion matrix in Figure 4 shows the TPR/FNR analysis of the classifier per true class plotted against the predicted class. In the distribution, Analysis, Backdoor, DOS, and Worms had high FNR for incorrectly classified attacks of 89.1%, 95.2%, 62.9%, and 65.0% respectively. These values indicate incorrect predictions with regards to these classes of attacks, with matrix indicating high FNR values; leading to the reduction in reliability and integrity of the model. The Exploits, Fuzzers, Generic, Normal, Reconnaissance, and shellcode had TPR values of 82.7%, 84.2%, 96.7%, 100%, 79%, and 73.8% respectively, indicating the percentages of the correctly classified points in each class. The adopted techniques successfully differentiated normal traffic but predicted few attacks as normal.

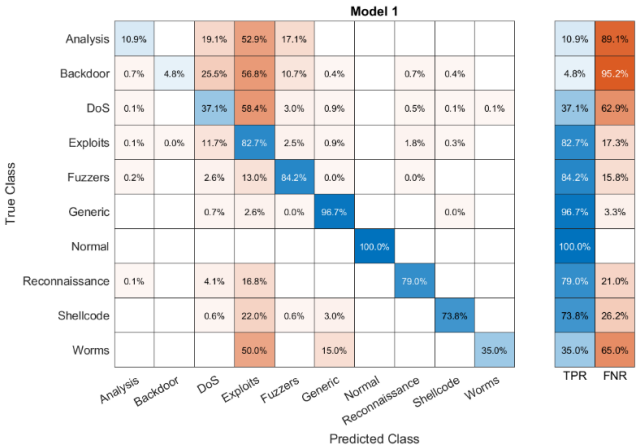

Figure 4: Confusion matrix showing TPR, FNR on UNSW-NB15 Dataset

Science World Journal Vol. 17(No 1) 2022
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

Figure 5 shows the results of Positive Predictive Values (PPV) and False Discovery Rates (FDR) which helps to discover the proportion of correctly and incorrectly classified observations per predicted class. Analysis, Backdoor, Exploits, Fuzzers, Generic, Normal, reconnaissance, Shellcode and worms had PPV values of 72.5%, 86.7%, 64.0%, 90.2%, 99.3%, 100%, 92%, 87.1% and 87.5% respectively. This implies that the classes have a high proportion of correctly classified attacks. However, the DOS had the highest FDR value of 58.8% which indicates that the form of attack was predicted under different attack classes.

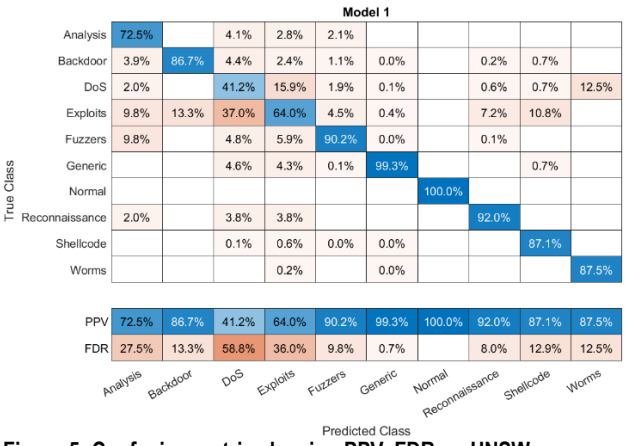

Figure 5: Confusion matrix showing PPV, FDR on UNSW-NB15 Dataset

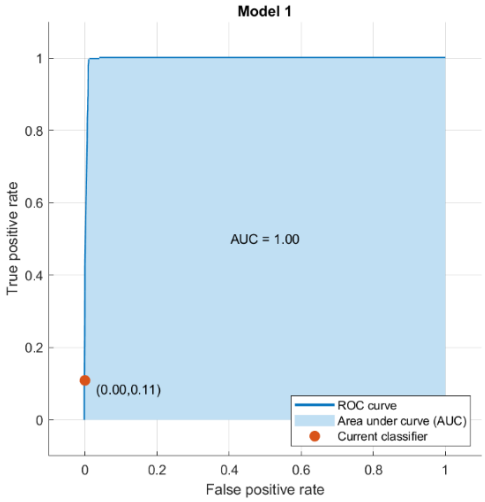

Figure 6: ROC Curve showing TPR and FDR on UNSW-NB15 Dataset

Figure 6 displays the ROC curve showing TPR, FPR, and AUC. The AUC values of 1.00 indicate 100% ability of the classifier to distinguish between the different classes of attacks which is an excellent classification ability.

Analysis of the Local Dataset

Executing the classification model with the locally generated dataset resulted in an accuracy of 99.4% with training times of 72.345sec. The Ensemble classifier was used for classification with the following model type settings; Preset: Optimizable Ensemble, Learner Type; Decision tree. The result of the experiments conducted is displayed in the confusion matrix and ROC curve as shown in figures 7 to 10.

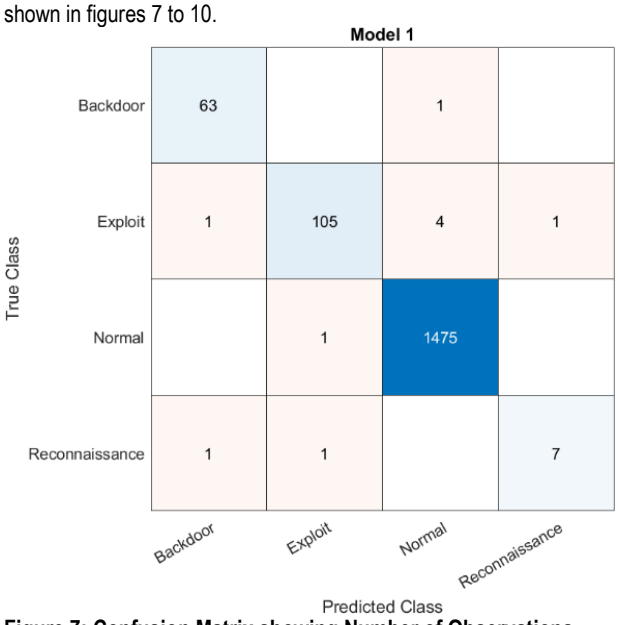

Figure 7: Confusion Matrix showing Number of Observations on Local Dataset

The correctly predicted classes for Backdoor, Exploits, normal traffic and reconnaissance are 63, 105, 1475, and 7 respectively. A total of 6and 4 FPR and FNR were recorded which indicates a low rate of wrongly classified categories of attacks to other classes.

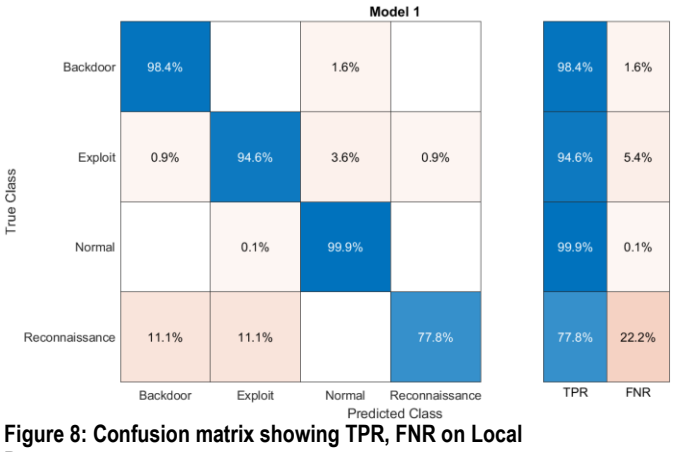

Figure 8: Confusion matrix showing TPR, FNR on Local Dataset

The confusion matrix in Figure 8 shows the TPR/FNR analysis of the classifier per true class plotted against the predicted class. In the distribution, all classes had a high TPR value of 98.4%, 94.6%, 99.9%, and 77.8% for Backdoor, Exploits, Normal, and reconnaissance attacks respectively. The TPR values show the percentage of the correctly classified points in each class. The adopted techniques successfully predicted a high percentage of the attacks to be in their true categories. In the distribution, low FNR for incorrectly classified attacks of 22.2%, 0.1%, 5.4% and 1.6% respectively.

Science World Journal Vol. 17(No 1) 2022
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

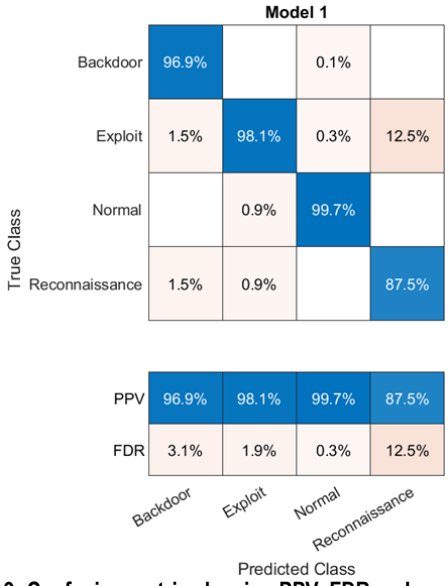

Figure 9: Confusion matrix showing PPV, FDR on Local Dataset

Figure 9 shows the results of PPV and FDR to discover the proportion of correctly and incorrectly classified observations per predicted class. PPV values of 96.9%, 98.1%, 99.7% and 87.5% respectively were obtained. This implies the classes have a good proportion of correctly classified attacks. With low FDR values of 3.1%, 1.9%, 0.3%, and 12.5% respectively, this implies forms of attack predicted wrongly under other attacks were very low.

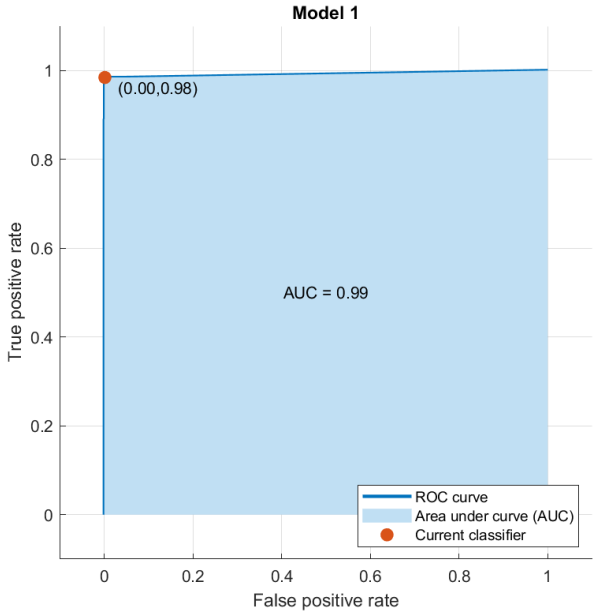

Figure 10: ROC Curve showing TPR and FDR on Local Dataset

Figure 10 displays the ROC curve showing TPR, FPR, and AUC. The AUC value of 0.99 indicates a 99% ability of the classifier to distinguish between the different classes of attacks indicating a good classification ability.

Comparison with A Selected State-of-Art Technique

Comparing the overall results of the two datasets, it is observed that both experiments recorded false positive and negative rates which indicate the need for an improvement in the area of false predictions. Table 1 shows the comparison of the adopted technique against selected state-of-the-art techniques(Aleesa et al., 2021), from their experimental setup, the authors did not state clearly how the benchmark datasets were setup. Also, most studies did not report false predictions recorded in their study. The benchmark set used for the study is the first two sets of UNSW-NB15 datasets which gave a total instance of 1,400,002.

Table 1: Comparison of the Proposed Model against selected state-of-art models.

	Local Datasets Results (%)						Public Dataset Results (%)					
Model Type	**ACC**	FNR	TPR	FDR	PPV	ROC	**ACC**	FNR	TPR	FDR	PPV	ROC
Ensemble	**99.4**	7.34	74.14	4.45	95.55	99	**99.1**	39.58	60.42	17.95	82.05	100
ANN(Aleesa et al., 2021)	**92.1**	53.35	46.65	8.9	91.1	63	**99.5**	47.11	52.89	20.41	59.59	100

In both experiments conducted, the ensemble gave a better false prediction and true positives compared to ANN (Aleesa et al., 2021). It is also clear from the results that there was a better performance with AUC of 99% and 100% as against 63% and 100% respectively.

CONCLUSION

This paper proposed an Ensemble model to predict uncertainties in cyberspace using MATLAB. The results of the model showed its efficiency in detecting abnormalities in NIDS datasets compared to other techniques with an overall accuracy of 99.4% for the penetration dataset and 99.1% for the local and UNSW-NB15 datasets. The results were generally good except for the fact that

there were False Positives (FP) and False Negatives (FN) and some attack categories were incorrectly classified. This could be used in a real-life scenario with firewalls to secure networks and reduce the impact of uncertainties experienced in cyberspace. Results obtained showed an increased percentage of true positive values for the ensemble model, this shows the model has a more robust prediction on the two datasets.

Future work is recommended on the improvement of the resulting confusion matrix by deploying techniques that can reduce the number of FPs and FNs. Similarly, classification learners should provide support for more instances to be experimented on at a time.

Acknowledgment

The authors will like to appreciate the support of the Nigerian Defence Academy for the opportunity and support given throughout this study.

REFERENCES

AL-DAWERI, M. S., ABDULLAH, S., ZAINOL ARIFFIN, K. A. & EFENDY, M. F. 2020. An Analysis of the KDD99 and UNSW-NB15 Datasets for the Intrusion Detection System. *Symmetry*, 1-32.

AL-ZEWAIRI, M., ALMAJALI, S. & AWAJAN, A. Experimental Evaluation of a Multi-Layer Feed-Forward Artificial Neural Network Classifier for Network Intrusion Detection System. International Conference on New Trends in Computing Sciences, 2017. 167-172.

ALEESA, A., THANOUN, M. Y., MOHAMMED, A. & SAHAR, N. 2021. Deep-Intrusion Detection Sysytem With Enhanced UNSW-NB15 Dataset based on Deep Learning Techniques. *Journal of Engineering Science and Technology,* 16**,** 711-727.

ALJUMAH, A. & AHAMAD, T. 2016. A Novel Approach for Detecting DDoS using Artificial Neural Networks. *International Journal of Computer Science and Network Security (IJCSNS),* 16**,** 132-138.

AWUJOOLA, O. J., OGWUELEKA, F. N., IRHEBHUDE, M. E. & MISRA, S. 2021. Wrapper Based Approach for Network Intrusion Detection Model with Combination of Dual Filtering Technique of Resample and SMOTE. *In:* MISRA, S. & KUMAR TYAGI, A. (eds.) *Artificial Intelligence for Cyber Security: Methods, Issues and Possible Horizons or Opportunities.* Cham: Springer International Publishing.

BRAMER, M. 2013. Ensemble Classification. *Principles of Data Mining.* London: Springer London.

ELKASSABI, H., ASHOUR, M. & ZAKI, F. 2020. Enhancing Classification of Network Intrusion Attacks using Feature Reduction. *International Journal of Scientific & Engineering Research (IJSER),* 11**,** 1643-1650.

FAKER, O. & DOGDU, E. Intrusion Detection Using Big Data and Deep Learning Techniques. ACM Southeast Conference (ACMSE), 2019. Association for Computing Machinery.

FREUND, Y. 1990. Boosting a weak algorithm by majority. *Information and Computation,* 12**,** 2.

IDHAMMAD, M., AFDEL, K. & BELOUCH, M. 2017. DoS Detection Method based on Artificial Neural Networks. *International Journal of Advanced Computer Science and Applications (IJACSA),* 8**,** 465-471.

KANIMOZHI, V. & JACOB, P. 2019. UNSW-NB15 Dataset Feature Selection and Network Intrusion Detection using Deep Learning. *International Journal of Recent Technology and Engineering (IJRTE),* 7**,** 433-446.

MAZINI, M., SHIRAZI, B. & MAHDAVI, I. 2018. Anomaly Network-Based Intrusion Detection System Using a Reliable Hybrid Artificial Bee Colony and AdaBoost Algorithms. *Journal of King Saud University - Computer and Information Sciences,* 31**,** 541-553.

MOUSTAFA, N. & SLAY, J. 2016. The Evaluation of Network Anomaly Detection Systems: Statistical Analysis of the UNSW-NB15 Dataset and the Comparison of KDD99. *Information Security Journal*, 1-14.

NAWIR, M., AMIR, A., YAAKOB, N. & LYNN, O. B. 2018. Multi-Classification of UNSW-NB15 Dataset for Network Anomaly Detection System. *Journal of Theoretical and Applied Information Technology,* 96**,** 5094-5104.

RASHID, M., KAMRUZZAMAN, J., IMAM, T., WIBOWO, S. & GORDON, S. 2022. A tree-based stacking ensemble technique with feature selection for network intrusion detection. *Applied Intelligence*.

SCHAPIRE, R. E. 1990. The strength of weak learnability. *Machine learning,* 5**,** 197-227.

SERAPHIM, I., PALIT, S., SRIVASTAVA, K. & POOVAMMAL, E. 2019. Implementation of Machine Learning Techniques Applied to the Network Intrusion Detection System. *International Journal of Engineering and Advanced Technology (IJEAT),* 8**,** 2721-2726.

SHARMA, N. & YADAV, N. S. Ensemble Learning based Classification of UNSW-NB15 dataset using Exploratory Data Analysis. 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 3-4 Sept. 2021 2021. 1-7.

SRINIVASAN, S., KETHA, S., RAVI, V., SOMAN, A. & KP, S. 2020. Towards Evaluating the Robustness of Deep Intrusion Detection Models in Adversarial Environment. *Security in Computing and Communications.*

YANG, L. 2011. Classifiers Selection for Ensemble Learning Based on Accuracy and Diversity. *Procedia Engineering,* 15**,** 4266 – 4270.

YUDHANA, A., RIADI, I. & RIDHO, F. 2018. DDoS Classification Using Neural Network and Naïve Bayes Methods for Network Forensics. *International Journal of Advanced Computer Science and Applications (IJACSA),* 9**,** 177-183.

ZHOU, Z.-H., WU, J. & TANG, W. 2002. Ensembling neural networks: Many could be better than all. *Artificial Intelligence,* 137**,** 239-263.