

Experimental analysis of intrusion detection systems using machine learning algorithms and artificial neural networks

Ademola Abdulkareem¹, Tobiloba Emmanuel Somefun¹, Adesina Lambe Mutalub²,
Adewale Adeyinka¹

¹Department of Electrical and Information Engineering, Covenant University, Ota, Nigeria

²Department of Electrical and Computer Engineering, Kwara State University, Kwara State, Nigeria

Article Info

Article history:

Received Oct 3, 2022

Revised Mar 15, 2023

Accepted Apr 3, 2023

Keywords:

Artificial neural

Ensemble classifier

Intrusion detection system

Machine learning

Networks attack

ABSTRACT

Since the invention of the internet for military and academic research purposes, it has evolved to meet the demands of the increasing number of users on the network, who have their scope beyond military and academics. As the scope of the network expanded maintaining its security became a matter of increasing importance. With various users and interconnections of more diversified networks, the internet needs to be maintained as securely as possible for the transmission of sensitive information to be one hundred per cent safe; several anomalies may intrude on private networks. Several research works have been released around network security and this research seeks to add to the already existing body of knowledge by expounding on these attacks, proffering efficient measures to detect network intrusions, and introducing an ensemble classifier: a combination of 3 different machine learning algorithms. An ensemble classifier is used for detecting remote to local (R2L) attacks, which showed the lowest level of accuracy when the network dataset is tested using single machine learning models but the ensemble classifier gives an overall efficiency of 99.8%.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Tobiloba Emmanuel Somefun

Department of Electrical and Information Engineering, Covenant University

Canaan Land, KM 10, Idiroko Road, P. M. B. 1023, Ota, Ogun State, Nigeria

Email: tobi.shomefun@covenantuniversity.edu.ng

1. INTRODUCTION

Access to the internet is very crucial to every business and individual in the 21st century [1], [2]. It is nearly impossible to compete in today's business world without staying connected to the world and customers. Staying connected to the internet is advantageous in the business world, but these advantages are not equipped to eliminate the accompanying threats, and it would be a disaster in this 21st century cyber-age and cyberspace if the power of a single click on the internet is ever underestimated [3], [4]. The possibility of these threats gave rise to the need for protective measures on the internet [5], [6]. Many confidential transactions occur every second. These exchanges on the web give an approach to unfrosted gatherings outside to obtain entrance into an organization's private organization and mess with the inside climate, data, assets, and structure. Network security helps us maintain the authorized access of data from hackers and authenticated data transfers, and we achieve the security of the network when a firewall is installed and turned ON.

With the rise in internet and network use [7], the need for security has become tantamount to user's convictions and interest to perform sensitive functions and activities on the internet or any cloud-based network system [8]–[10]. As the internet evolves, likewise the various malicious software hosted on the network and the attacks have become increasingly sophisticated [11]. In a 2017 report released by Symantec,

on internet security threat, it recorded over three billion zero-day assaults in 2016, this implied that the assaults were gaining popularity and becoming increasingly common unlike before [12]. The 2017 data breach statistics recorded around nine billion lost or hijacked information records since 2013. A Symantec report tracked down that the quantity of safety penetrate occurrences is rising rapidly [13]. Various malicious software that penetrates internal company networks have become more sophisticated, directly affecting the severity of attacks companies experience, even as security measures evolve with time [14]. Several reports have revealed that security breaches are consistently on the rise. Tactics of cybercriminals have begun to change with the times, and as some researchers would describe it, more ambitious [15], [16]. Previously these attackers targeted “smaller fish” like credit cards, bank customers, bank accounts, whereas these days, they target the banks themselves [17]–[20]. All these are possible because of the evolution of malicious software [21]–[23]. Malicious Software (Malware) is intentionally designed to take advantage of any compromise, or weakness however minute, in the firewall to gain access to the inside network.

A survey carried out by Kaspersky in 2013 revealed that 91% of companies had experienced at least one security threat from outside the company network, 35% of these companies encountered data leakage due to these attacks [24]. 61% of these companies were attacked by spam, while another 66% of the companies were affected by viruses, spyware, malware, worms, and other malicious programs. Even though the attack rate is this high, the discovery rate for malware and intrusions is still low [25], [26]. In Panda Lab’s 2015 annual report, the following discoveries were made; 34% of all malwares were produced in 2014. 65% of attacked systems were intruded on by Trojans, making Trojans the major contributor of security threats. This report concludes that despite the depth of research and development of network security infrastructure, online inform action will still be exploited by new forms of attack [27].

Cloud infrastructure utilizes integrated technologies, virtualization techniques, and it moves according to standard internet protocols, which may attract unauthorized users due to the weaknesses present in the cloud infrastructure. Distributed computing experiences different conventional assaults that include protocol spoofing, address resolution, internet protocol (IP) spoofing, flooding, distributed denial of service (DDoS), domain name system (DNS), poisoning, denial of service (DoS), and routing information protocol attack. A genuine model is the DoS assault on the fundamental Amazon Cloud framework that caused BitBucket.org, a site facilitated on Amazon web services (AWS), to stay inaccessible for a couple of hours [28]. Firewalls can be an effective method to protect a network from external attacks, but it is not applicable for internal attacks; therefore, an efficient intrusion detection system (IDS) should be fused with Cloud infrastructure to alleviate these attacks. In this study, authors seek to find out the cause for the attacks on the networks and investigate ways to identify and curb these attacks. Also, to discover and recommend better security measures for the protection of networks and network-based systems from security attacks/threats.

2. MATERIAL AND METHOD

In this study, the Network Security Laboratory-knowledge discovery in databases (NSL-KDD) dataset is used instead of the original KDD Cup 99 dataset, because it gives a good understanding of intrusion behaviors. Six processes were involved in the approach followed in this study which are data collection, data pre-processing, feature scaling, feature selection, model development, accuracy evaluation.

The NSL-KDD dataset, which comprises network packets with 42 attributes is used for data collection. The data is thereafter pre-processed into a suitable form to be utilized by the algorithm. Pre-processing involves cleaning the algorithm to remove duplicate and redundant entries. Every feature is transformed to a numerical value/feature by “one-Hot encoding,” which converts objects/string values into categorical data and is then converted to numerical data using label Encoder in-built in Python. To avoid features with large values that may weigh too much in the results and eventually lead to overfitting, the features must be scaled. After the conversion, the dataset is split into 4 different datasets, each representing the different attack categories. Attack categories are shown in Table 1. The attack categories are renamed as 0=normal, 1=DoS, 2=Probe, 3=R2L, 4=U2R.

Table 1. Attack types in NSL-KDD dataset

Attack category	Attack name
Denial of service (DoS)	Apache2, Smurf, Neptune, Back, Teardrop, Pod, Land, Mailbomb, Processtable, UDPstrom
Remote to local (R2L)	WareZMaster, Imap, Ftp_Write, Named, MultiHop, Phf, Spy, Sendmail, SntpGetAttack, SntpGuess, Worm, Xsnoop, Xlock
User to root (U2R) probe	Buffer_Overflow, Httpstuneel, Rootkit, LoadModule, Perl, Xterm, Ps, SQLattack, Satan, Saint, Ipsweep, Portsweep, Nmap, Mscan

StandardScaler() library is used to scale the data frames and ensure the standard deviation is 1. The univariate feature selection using analysis of variance (ANOVA) F-test (second percentile method) is first used, followed by the recursive feature elimination (RFE) method, to get the best features for each dataset. The formula for each classifier is already built-in to Python, so each attack dataset goes through all the different classification algorithms before producing results.

2.1. Decision tree classifier

A decision tree (DT) classifier is a popular machine learning algorithm used for both classification and regression tasks. It recursively partitions a dataset into subsets based on the most significant features, effectively creating a tree-like structure of decisions. These splits are determined by various criteria, with one common measure being Gini impurity, which quantifies the randomness or impurity in each subset. Subset is determined using (1).

$$\sum_{i=1}^c f_i(1 - f_i) \quad (1)$$

where f_i is the frequency of labels at a node, and c is the number of unique labels.

2.2. Support vector machine classifier

A support vector machine (SVM) classifier aims to find the optimal hyperplane that best separates different classes in the feature space. By maximizing the margin between data points and the hyperplane, SVM enhances its generalization performance, proving especially effective in high-dimensional spaces commonly encountered in image and text analysis. The hyperplane position is determined by support vectors, which are the data points closest to the decision boundary, playing a crucial role in defining the classification boundary accurately.

2.3. K-nearest neighbors algorithm

Unlike some others, the k-nearest neighbors (KNN) is non-parametric, which implies that it makes no assumptions about the underlying data. It can be used for both regression and classification problems but primarily for classification. This algorithm stores data such that when a new data entry is made, it quickly classifies it based on its similarity to already existing data points. Classification algorithm; given a query instance x_q to be classified, let x_1, \dots, x_k denote the k instances from the training examples.

$$\text{Return } f(x_q) \leftarrow \arg \max \sum_{i=1}^k \delta(v, f(x_i)) \text{ for the discrete-valued target function}$$

where $(a, b) = 1$ if $a=b$ and where $\delta(a, b) = 0$, otherwise. The weights of neighbors are taken into consideration relative to their distance to the query point such that:

$$f(x_q) \leftarrow \operatorname{argmax}_{v \in V} \sum_{i=1}^k w_i \delta(v, f(x_i)) \quad (2)$$

$$\text{where } w_i = \frac{1}{d(x_q, x_i)^2}.$$

2.4. Artificial neural network classifier

Artificial neural network (ANN) is a supervised machine learning (ML) algorithm that is based on the human brain. The advantage of using this algorithm is its performance ability in nonlinear modelling. Also, because of it is various layers, it provides a more accurate representation of the predictions. In developing this model, the dataset is fed into the model 5 times to make provisions for the system memory and improve the accuracy metric for each attack type.

2.5. Ensemble classifier

The dataset is run through the different classification algorithms that have been previously used. it goes through the DT, KNN, and SVM classifiers one after the other. This is also done to measure for an improved accuracy compared to the individual testing and training carried out on the dataset by each classification algorithm.

3. RESULT AND DISCUSSION

This section discusses the implementation of the machine learning algorithms discussed in section 2. Furthermore, it explains commonly used evaluation metrics for machine learning methods for IDS. The

general confusion matrix, which is used to visualize the performance of our supervised learning algorithms is shown in Table 2.

Table 2. Confusion matrix

Actual Class	Predicted class	
	Attack	Normal
Attack	True positive	False negative
Normal	False positive	True negative

3.1. DoS attack

After running our DoS attack dataset through this decision tree, SVM, and KNN classifiers, the results are shown in Tables 3 and 4. Table 3 shows the confusion matrix for DoS attacks, classified using the three stated classifiers algorithm, while Table 4 shows other metrics tested for by the classifiers. Metrics such as precision, recall, accuracy, and F-measure.

Table 3. Confusion matrix for three classifiers on DoS attack

DoS attack	Predicted attacks		Classifier	
Actual attacks	0	1		
	0	9,602	109	DT
	1	2,625	485	
		0	1	SVM
	0	9,677	34	
	1	3,578	3,882	KNN
		0	1	
	0	9,653	58	
	1	2,645	4,815	

Table 4. Evaluation metrics for three classifiers on DoS attack

Metrics	Precision	Recall	F-Measure	Support	Classifier
0	0.79	0.99	0.88	9,711	DT
1	0.98	0.65	0.78	7,460	
Accuracy	-	-	0.84	17,171	
Macro avg	0.88	0.82	0.83	17,171	
Weighted avg	0.87	0.84	0.83	17,171	SVM
0	0.73	1	0.84	9,711	
1	0.99	0.52	0.68	7,460	
Accuracy	-	-	0.79	17,171	
Macro avg	0.86	0.76	0.76	17,171	KNN
Weighted avg	0.84	0.79	0.77	17,171	
0	0.78	0.99	0.88	9,711	
1	0.99	0.65	0.78	7,460	
Accuracy	-	-	0.84	17,171	KNN
Macro avg	0.89	0.82	0.83	17,171	
Weighted avg	0.87	0.84	0.83	17,171	

The classifier resulted in 9,602 correctly predicted attacks from the 12,821 data entries/input. Only 485 out of the 3,110 standard entries were accurately predicted as regular attacks by this classifier. This is shown in Table 3. From this result, we can see that this decision tree classifier produces better attack predictions compared to typical network behavior. The accuracy of this method is 0.84 but can be improved on. This will be revealed in the results of the ANN and ensemble classifier.

After running the DoS attack dataset through the SVM classifier, which uses a subset of training points in the decision function. The accuracy of the classification algorithm is measured using the metrics recorded is 0.79. The KNN classifiers is a simple algorithm that stores and classifies cases based on similarity measures such as distance functions. The confusion matrix for this classifier shows a more robust prediction for the regular network behavior compared to that of the decision tree and strongly predicts the DoS attacks. The accuracy for this classifier is 0.84, like that of the decision tree classifier.

The results from the ANN classifier for the DoS attacks using the tensor flow framework in Python is a loss metric of 0.0602, and an accuracy of 0.975. The graphs Figure 1 compares the three classifiers with ANN. In Figure 1, all precision metrics for the DoS attacks are mapped out, and the DT algorithm has a more precise measure for DoS attacks.

The accuracy of predicting DoS attacks using all the different classification algorithms is compared in Figure 1, and it is evident that the ANN has higher accuracy. This can result from the deep neural networks utilized in developing the ANN model, unlike the other machine learning algorithms where the dataset is fed into the classifier only once.

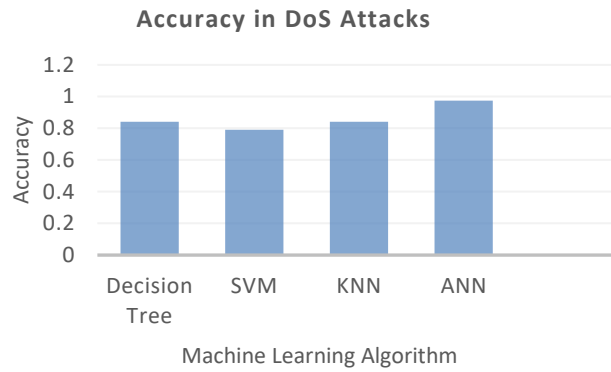


Figure 1. Accuracy in DoS attack

3.2. Probe attacks

These attacks are deliberately crafted so that the legitimate users of the network recognize the intrusion and report it. After reporting the attack, the attacker uses recognizable fingerprints to learn more about the network capabilities. After running the dataset through the decision tree, SVM, and KNN classifiers, the results from the confusion matrix are shown in Table 5. It shows the ability of the classifier to predict attacks accurately. The results signify that the decision tree classifier may not be the best for predicting probe attacks. Inasmuch as the false negative and false positives are less than the true negative and true positive, the values are still relatively large. The accuracy of this classification algorithm is found to be 84% as shown in Table 6.

Table 5. Confusion matrix for three classifiers on probe

Probe attack	Predicted attacks		Classifier
Actual attacks	0	2	DT
	0	8,709 1,002	
1	944	1,477	SVM
	0	2	
0	9,074	637	KNN
	1	958 1,463	
1	0	2	KNN
	0	9,107 604	
1	943	1,478	

Table 6. Evaluation metrics for the three classifiers on probe

Metrics	Precision	Recall	F-Measure	Support	Classifier
0	0.9	0.9	0.9	9,711	DT
1	0.6	0.61	0.6	7,460	
Accuracy	-	-	0.84	17,171	
Macro avg	0.75	0.75	0.75	17,171	
Weighted avg	0.84	0.84	0.84	17,171	
0	0.9	0.93	0.92	9,711	SVM
1	0.7	0.6	0.65	7,460	
Accuracy	-	-	0.87	17,171	
Macro avg	0.8	0.77	0.78	17,171	
Weighted avg	0.86	0.87	0.86	17,171	
0	0.91	0.94	0.92	9,711	KNN
1	0.71	0.61	0.66	7,460	
Accuracy	-	-	0.87	17,171	
Macro avg	0.81	0.77	0.79	17,171	
Weighted avg	0.87	0.87	0.87	17,171	

SVM can be used for regression and classification. Since this is a classification problem, it is used here for classification. It works by finding an optimal boundary between two outputs. Accuracy of this classifier is 87%.

The results gotten from the KNN classifier are shown in Tables 5 and 6. There is a significantly high prediction possibility, evident in the true negative and true positive values. The accuracy of this classification algorithm is measured to be 87% and given as the output of the code in Python.

The ANN classifier evaluated in the tensor flow framework of the Python IDE gives an accuracy of 88.7%, with a loss measure of 0.321. The loss in this classification algorithm is high. The accuracy measure is not as high as expected because information security must be optimal enough to predict over 90% of attacks. From Figure 2, it is clear that the ANN classifier has the highest accuracy, which the presence of more layers can explain unlike the single layers of the other machine learning algorithms.

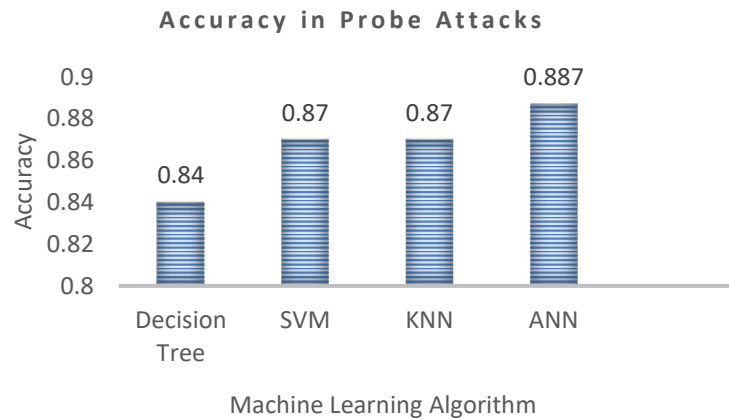


Figure 2. Accuracy in probe attacks

3.3. R2L attack

The R2L attack type represents a scenario where a user without remote network access attempts to send packets to gain unauthorized entry. In the context of our analysis, the decision tree classifier's performance in detecting these R2L attacks is depicted in the confusion matrix displayed in Table 7. This matrix reveals a remarkably high prediction rate, underscoring the effectiveness of the decision tree model in identifying and mitigating such intrusion attempts.

Table 7. Confusion matrix for three classifiers on probe

Probe attack	Predicted attacks		Classifier
Actual attacks	0	3	DT
	9,649	62	
	1	2,560	
	0	0	SVM
	9,711	0	
	3	2,885	
	0	3	KNN
	9,710	1	
	3	2,885	

The accuracy measure of this method is gotten to be 79% as shown in Table 8. This is not a very high accuracy for internet security, so we will use other classification algorithms to decide on the model with the highest accuracy. The results outputted from the code for this classifier show us an accuracy level of 77%. This accuracy level is not good enough for network security purposes, so other classification algorithms and ANN are used to analyses the accuracy levels.

The accuracy of this classification algorithm is also 77% which is still not good enough for network security. So far, we have seen that machine learning algorithms are not the best for predicting R2L attacks. The ANN classifier gives an output of 0.9998 and a loss of 0.003. This accuracy level is very efficient for a network security prediction model. Figure 3 is a graphical representation of the different classification

algorithms used to analyses the R2L dataset. The ANN classifier produces a more robust accuracy, unlike the machine learning algorithms.

Table 8. Evaluation metrics for the three classifiers on probe

Metrics	Precision	Recall	F-Measure	Support	Classifier
0	0.79	0.99	0.88	9,711	DT
1	0.84	0.11	0.2	7,460	
Accuracy	-	-	0.84	17,171	
Macro avg	0.82	0.55	0.54	17,171	SVM
Weighted avg	0.8	0.79	0.72	17,171	
0	0.79	0.99	0.88	9,711	
1	0.84	0.11	0.2	7,460	KNN
Accuracy	-	-	0.84	17,171	
Macro avg	0.82	0.55	0.54	17,171	
Weighted avg	0.8	0.79	0.72	17,171	ANN
0	0.77	1	0.87	9,711	
1	0	0	0	7,460	
Accuracy	-	-	0.77	17,171	ANN
Macro avg	0.39	0.5	0.44	17,171	
Weighted avg	0.59	0.77	0.67	17,171	

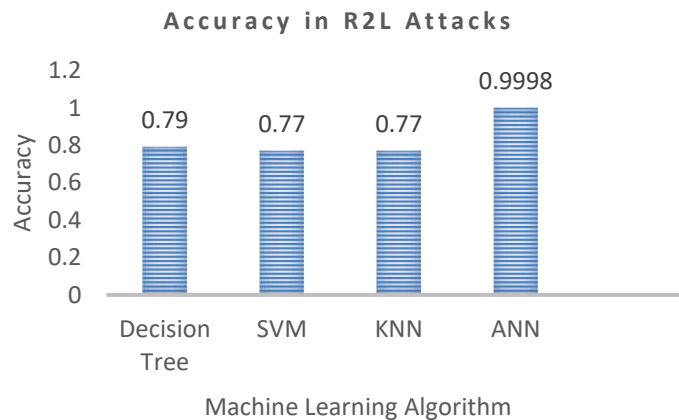


Figure 3. Accuracy in R2L attacks

3.4. U2R attacks

User 2 Root (U2R) attack is the illegal access of the root of a network by a local user who has only been granted access to the leading network, not the network's backend. With the considered three classifiers in this study, their true positive, true negative, false positive, and false negative are shown in Table 9 while the other metrics are given in Table 10. The accuracy output of this classification metric is very high at 99%, making it very efficient and appropriate for predicting network attacks.

As shown in Figure 4, the accuracy for the U2R attacks using the SVM classifier has an extremely high accuracy of 99%. This signifies that the SVM classifier efficiently predicts future U2R attacks on a network. Also, the accuracy of this KNN classifier is 99%, meaning it would be very efficient in predicting attacks and protecting the network from intrusion. The results from the ANN classification produced an output prediction value of 99.69%.

3.5. Ensemble classifier

The ensemble classifier is a combination of the various classifier previously used. This is experimented upon the dataset to determine the accuracy of identifying attacks. The ensemble classification was carried out on the R2L attack to see if there will be an increase in it is 77% accuracy, which was obtained from the other independent machine learning algorithms. The output from the ensemble classifier outputted a whopping accuracy of 99.98%.

From the results output, which has been visualized in the Table 8 and Figure 5, it is clear that using the ANN classifier is the most accurate way to predict network attacks and intrusions. The ANN classifier produces results that are close in metric to the ensemble classification, i.e., the combination of the various machine learning algorithms.

Table 9. Confusion matrix for the classifier on U2R

Probe attack	Predicted attacks			Classifier
Actual attacks	0	5		DT
	0	9,706	5	
	4	52	15	SVM
	0	0	4	
	0	9,711	0	KNN
	4	67	0	
	0	0	4	
	0	9,709	2	
	4	60	7	

Table 10. Evaluation metrics for the three classifiers on U2R

Metrics	Precision	Recall	F-Measure	Support	Classifier
0	0.99	1	1	9,711	Decision Tree
1	0.75	0.22	0.34	67	
Accuracy	-	-	0.99	8,778	
Macro avg	0.87	0.61	0.67	9,778	
Weighted avg	0.99	0.99	0.99	9,778	SVM
0	0.99	1	1	9,711	
1	0	0	0	67	
Accuracy	-	-	0.99	8,778	
Macro avg	0.5	0.5	0.5	9,778	KNN
Weighted avg	0.99	0.99	0.99	9,778	
0	0.99	1	1	9,711	
1	0.78	0.1	0.18	67	
Accuracy	-	-	0.99	8,778	
Macro avg	0.89	0.55	0.59	9,778	
Weighted avg	0.99	0.99	0.99	9,778	

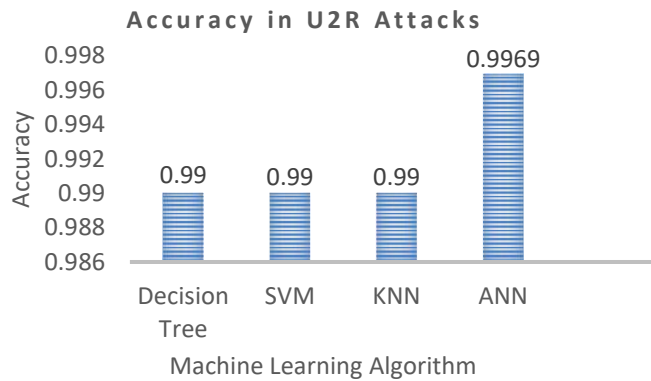


Figure 4. Accuracy in U2R attacks

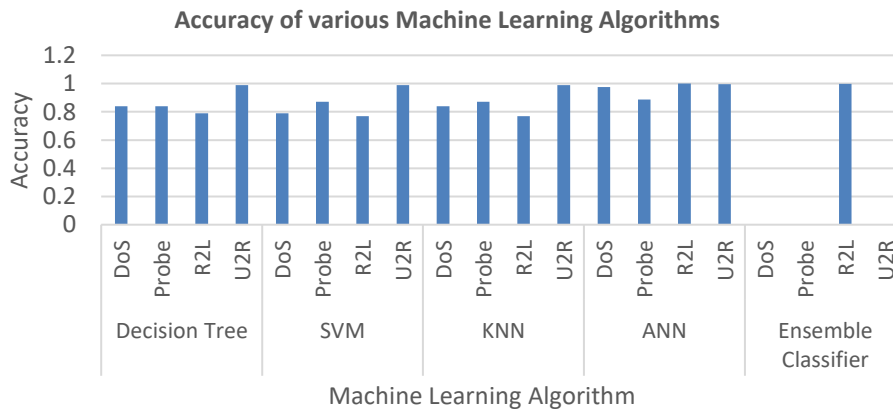


Figure 5. Accuracy in U2R attacks

4. CONCLUSION

In this study, various network intrusions were analyzed using several machine learning algorithms as classifiers. This was to see how accurately, and intelligently various machine learning algorithms detect network intrusions when encountered in a system. These experiments were carried out to analyses the NSL-KDD dataset, which revealed that the dataset is ideal for comparing intrusion detection models. 99% accuracy was obtained on some of the intrusion detection models developed. The experiments have demonstrated that there is no single machine learning algorithm that can efficiently handle all types of attacks, but the models can be trained to give efficiencies up to 99.98% which will tremendously predict and prevent attacks from flooding the network.

ACKNOWLEDGEMENTS

The authors will like to appreciate Covenant University for her financial support.





REFERENCES

- [1] M. Bala and D. Verma, "A critical review of digital marketing," *A Critical Review of Digital Marketing. International Journal of Management, IT and Engineering*, vol. 8, no. 10, pp. 321–339, 2018.
- [2] M. Fahlevi, M. Saparudin, S. Maemunah, D. Irma, and M. Ekhsan, "Cybercrime business digital in Indonesia," *E3S Web of Conferences*, vol. 125, Oct. 2019, doi: 10.1051/e3sconf/201912521001.
- [3] N. M. Sambaluk, *Myths and realities of cyber warfare*. ABC-CLIO, LLC, 2020.
- [4] H. Kim, H. Kwon, and K. K. Kim, "Modified cyber kill chain model for multimedia service environments," *Multimedia Tools and Applications*, vol. 78, no. 3, pp. 3153–3170, Feb. 2019, doi: 10.1007/s11042-018-5897-5.
- [5] L. Tawalbeh, F. Muheidat, M. Tawalbeh, and M. Quwaider, "IoT privacy and security: challenges and solutions," *Applied Sciences*, vol. 10, no. 12, Jun. 2020, doi: 10.3390/app10124102.
- [6] M. Z. Gunduz and R. Das, "Cyber-security on smart grid: Threats and potential solutions," *Computer Networks*, vol. 169, Mar. 2020, doi: 10.1016/j.comnet.2019.107094.
- [7] A. A. Adewale, A. S. Ibdunni, A. A. Atayero, S. N. John, O. Okesola, and R. R. Ominiabohs, "Nigeria's preparedness for internet of everything: A survey dataset from the work-force population," *Data in Brief*, vol. 23, Apr. 2019, doi: 10.1016/j.dib.2019.103807.
- [8] J. Li, "Cybercrime in the Philippines: A case study of national security," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 11, pp. 4224–4231, 2021.
- [9] F. Adeoye, "Issues in internet regulation in Nigeria: the need to promulgate a befitting legislation," *SSRN Electronic Journal*, 2020, doi: 10.2139/ssrn.3773010.
- [10] Z. Zhang *et al.*, "An overview of security support in named data networking," *IEEE Communications Magazine*, vol. 56, no. 11, pp. 62–68, Nov. 2018, doi: 10.1109/MCOM.2018.1701147.
- [11] O. Osemwegie, K. Okokpujie, N. Nkordeh, S. John, and A. A. Adeyinka, "On issues, strategies and solutions for computer security and disaster recovery in online start-ups," *International Journal of Applied Engineering Research*, vol. 12, no. 19, pp. 8009–8015, 2017.
- [12] C. S. Teoh and A. K. Mahmood, "National cyber security strategies for digital economy," in *2017 International Conference on Research and Innovation in Information Systems (ICRIIS)*, Jul. 2017, pp. 1–6, doi: 10.1109/ICRIIS.2017.8002519.
- [13] H. Hammouchi, O. Cherqi, G. Mezzour, M. Ghogho, and M. El Koutbi, "Digging deeper into data breaches: an exploratory data analysis of hacking breaches over time," *Procedia Computer Science*, vol. 151, pp. 1004–1009, 2019, doi: 10.1016/j.procs.2019.04.141.
- [14] L. Y. Connolly and D. S. Wall, "The rise of crypto-ransomware in a changing cybercrime landscape: Taxonomising countermeasures," *Computers and Security*, vol. 87, Nov. 2019, doi: 10.1016/j.cose.2019.101568.
- [15] A. Ayodele, J. K. Oyediji, and H. O. Badmos, "Social construction of internet fraud as innovation among youths in Nigeria," *International Journal of Cybersecurity Intelligence and Cybercrime*, Mar. 2022, doi: 10.52306/BUVC2778.
- [16] D. N. Jones, E. Padilla, S. R. Curtis, and C. Kiekintveld, "Network discovery and scanning strategies and the Dark Triad," *Computers in Human Behavior*, vol. 122, Sep. 2021, doi: 10.1016/j.chb.2021.106799.
- [17] D. Airehrour, N. V. Nair, and S. Madanian, "Social engineering attacks and countermeasures in the New Zealand banking system: advancing a user-reflective mitigation model," *Information*, vol. 9, no. 5, May 2018, doi: 10.3390/info9050110.
- [18] T. Kellermann and R. Murphy, "Modern bank heists 3.0," *Annual "Modern Bank Heists." VMware Carbon Black*, 2020.
- [19] M. Botacin, A. Kalysch, and A. Grégio, "The internet banking [in] security spiral," in *Proceedings of the 14th International Conference on Availability, Reliability and Security*, Aug. 2019, pp. 1–10, doi: 10.1145/3339252.3340103.
- [20] M. Komar, V. Dorosh, G. Hladyi, and A. Sachenko, "Deep neural network for detection of cyber attacks," in *2018 IEEE First International Conference on System Analysis and Intelligent Computing (SAIC)*, Oct. 2018, pp. 1–4, doi: 10.1109/SAIC.2018.8516753.
- [21] A. Mpanti, S. D. Nikolopoulos, and I. Polenakis, "Malicious software detection utilizing temporal-graphs," in *Proceedings of the 20th International Conference on Computer Systems and Technologies*, Jun. 2019, pp. 49–55, doi: 10.1145/3345252.3345269.
- [22] M. N. Alenezi, H. K. Alabdulrazzaq, A. A. Alshaher, and M. M. Alkharang, "Evolution of malware threats and techniques: a review," *International Journal of Communication Networks and Information Security (IJCNIS)*, vol. 12, no. 3, Apr. 2022, doi: 10.17762/ijcnis.v12i3.4723.
- [23] H. Sultan, A. Khaliq, S. I. Alam, and S. Tanweer, "A survey on ransomware: evolution, growth, and impact," *International Journal of Advanced Research in Computer Science*, vol. 9, no. 2, 2018.
- [24] G. Tsochev, R. Trifonov, O. Nakov, S. Manolov, and G. Pavlova, "Cyber security: threats and challenges," in *2020 International Conference Automatics and Informatics (ICAI)*, Oct. 2020, pp. 1–6, doi: 10.1109/ICAI50593.2020.9311369.
- [25] Z. Lin, Y. Shi, and Z. Xue, "IDSGAN: generative adversarial networks for attack generation against intrusion detection," in *Advances in Knowledge Discovery and Data Mining*, Springer International Publishing, 2022, pp. 79–91.
- [26] N. Martins, J. M. Cruz, T. Cruz, and P. H. Abreu, "Adversarial machine learning applied to intrusion and malware scenarios: a systematic review," *IEEE Access*, vol. 8, pp. 35403–35419, 2020, doi: 10.1109/ACCESS.2020.2974752.





- [27] F. Salo, M. Injadat, A. B. Nassif, A. Shami, and A. Essex, "Data mining techniques in intrusion detection systems: a systematic literature review," *IEEE Access*, vol. 6, pp. 56046–56058, 2018, doi: 10.1109/ACCESS.2018.2872784.
- [28] H. Banafar and S. Sharma, "Secure cloud environment using hidden markov model and rule based generation," *International Journal of Computer Science and Information Technologies*, vol. 5, no. 3, pp. 4808–4841, 2014.

BIOGRAPHIES OF AUTHORS







Ademola Abdulkareem     is a renowned researcher and associate professor of electrical engineering in Covenant University. His areas of specialization are electrical and electronics engineering, power systems quality analysis and control and energy. His research interests are optimization of power system network and security, power system investment planning, high voltage engineering, smart power distribution system and intelligent building and energy management and renewable energy. He can be contacted at email: ademola.abdulkareem@covenantuniversity.edu.ng.







Tobiloba Emmanuel Somefun     is a Ph.D. holder from Covenant University in electrical and electronics engineering. His research interest covers energy management, power system analysis, and data analysis. He can be contacted at email: tobi.somefun@covenantuniversity.edu.ng.



Adesina Lambe Mutalub     currently works as a senior lecturer at the Electrical and Computer Engineering Department, Faculty of Engineering and Technology, Kwara State University, Malete, Nigeria. L. M. Adesina does research in engineering education and electrical engineering. He can be contacted at email: lambe.adesina@kwasu.edu.ng.



Adewale Adeyinka     obtained his Ph.D. in information and communication engineering at Covenant University. His research interests are quality of service (QoS) in mobile or wireless communication; computer network security; artificial intelligence application; digital signal processing; software design and programming; electric power distribution; electrical installations; and Seismic or geophysical data acquisition operations. He can be contacted at email: adeyinka.adewale@covenantuniversity.edu.ng.