Intrusion Detection using Associative Rule and Support Vector Machine

Fadekemi A. Adetoye
Department of computer Science,
School of Computing
Federal University of Technology,
Akure, Ondo State, Nigeria

ABSTRACT
In this contemporary technology-motivated era, shielding our private information from being accessed by unauthorized users is becoming more intricate, vastly confidential information are becoming more accessible by public databases, because we are more interconnected than ever. Thus, our information is available for almost anyone to filter due to this interconnectivity, and this creates a pessimistic mindset that the use of technology is hazardous, unpredictable and highly unprotective because virtually anyone can access one's private information for an outlay. The weaknesses discovered from the previous work are the key motivation for this research work. These includes: The work done on Network Intrusion Detection using Association Rules which generated an incomprehensive set of attack rules due to the small percentage of KDD'99 data set used for training set, proposed wrapper method for feature selection in multiple class data set using a sequential backward elimination method which is more computationally expensive and time consuming, and Development of a Denial of Service attack detection using machine learning technique in which the Significant features of data set were not extracted, and the extraction was done using only one extraction technique which results in high level of FAR (False Alarm Rate) due to poor detection of attacks. This research makes use of NSL-KDD and UNSW-NB15 data set, with filter and wrapper method as the feature selection techniques. In addition, an intrusion detection model was developed based association rule and support vector machine and performance of the model was evaluated.

General Terms
Rule based Intrusion Detection and non-rule-based intrusion detection

Keywords
Intrusion detection, Machine learning, security, Data analysis, Data Science.

1. INTRODUCTION
In this modern technology-driven age, protecting our personal information from being accessed by unauthorized users is becoming more difficult. Highly classified details are becoming more available to public databases, because we are more interconnected than ever [1]. Thus, our data is available for almost anyone to sift through due to this interconnectivity, and this creates a negative mindset that the use of technology is dangerous, unreliable and highly unprotective because practically anyone can access one's private information for a price [2]. Although, technology continues to promise to ease our daily lives; however, there are dangers of using technology is the threat of cybercrimes [3].

Information security is a matter of serious worldwide concern as the incredible development in connectivity and accessibility to the internet has generated tremendous security threat to information system worldwide.

An intrusion is defined as a set of activities that attempt to compromise the confidentiality, integrity or availability of resources. This includes a deliberate unauthorized attempt to access information, manipulate information or render a system unreliable or unusable [4]. An attacker can gain illegal access to a system by fooling an authorized user into providing information that can be used to break into the system, he can as well deliver a software which is actually a Trojan horse containing malicious code to a system user which enables attacker to gain access into the system [5]. An intrusion is defined as a set of activities that attempt to compromise the confidentiality, integrity or availability of resources. This includes a deliberate unauthorised attempt to access information, manipulate information or render a system unreliable or unusable [6]. An attacker can gain illegal access to a system by fooling an authorized user into providing information that can be used to break into the system; he can as well deliver a software which is actually a Trojan horse containing malicious code to a system user which enables attacker to gain access into the system [7].

Intrusion detection is defined as identifying unauthorized use, misuse and abuse of computer systems by both inside and outside intruders. The main task of an intrusion detection system (IDS) is to defend a computer system or computer network by detecting hostile attacks on a network system or host device, monitoring the events occurring in a computer system or network and analyzing them for signs of intrusions [8]. That compromise the integrity, confidentiality, and availability of information resources.

Security incidence resulting from attempted attacks violate the CIA (Confidentiality, Integrity and Availability) triads of computer security.

The National Institute of Standards and Technology (NIST) defines intrusion as an attempt to compromise CIA or to bypass the security mechanisms of a computer network [9]. The specific objectives of this research are; to extract relevant features or attributes of the NSL-KDD data set and UNSW-NB15 data set using Filter, and Wrapper methods, Design Intrusion Detection Model based on Association Rule and Support Vector Machine and Implement the model developed.

A common thought-provoking problem most scientists face is how to select the most approving/best conceivable set of features, as not all features are applicable and have impact on classification performance. In numerous situations, non-applicable features can impact the classification accuracy and cause slow training and testing process [10]. The key motivation for this research work includes:

The work done by [11] on Network Intrusion Detection using Association Rules which generated an incomprehensive set of attack rules due to the small percentage of KDD'99 data set used for training set.

The work done by [12] on proposed wrapper method for feature selection in multiple class data set, which uses a sequential backward elimination method which is more computationally expensive and time consuming.

The work done by [13] on Denial of Service attack detection using machine learning techniques in which the Significant features of data set were not extracted or extraction done using an extraction technique and High level of FAR (False Alarm Rate) due to poor detection of attacks.

In [14], a Network Intrusion Detection using Association Rule was developed. A technique to generate rules that detects attacks in network audit data using association rules algorithm on KDD' 99 data set was developed. In [12], a wrapper method for feature selection in multiple class data set was proposed, which describes a new wrapper method called IAFN-FS (Incremental ANOVA and Functional Networks-Feature Selection) as described in its version for dealing with multiclass problems. [13] developed denial of service attacks detection using machine learning techniques. The system minimized cases of denial of service attack and Bayesian methods. However, the system did not classify some attacks that used more than three rules in the database. [14], proposed a system which uses multi-layered Perception (MLP) architecture. The system detects attack and classifier into six groups. Author points out the issue of obtaining irrelevant output and suggested more research work in the future to obtain relevant output that could solve the weakness of the work. [15] presented a paper on Understanding Modern Intrusion Detection Systems: A Survey: This paper presents a general overview of IDSs, the way they are classified, and the different algorithms used to detect anomalous activities. [16] proposed a hybrid method as a feature selection, based on the central points of attribute values and an Association Rule Mining algorithm to decrease the FAR. [17] presents Network Anomalies Classifier (NAC) that uses machine learning technologies to automatically classify activities detected by a packet header-based anomaly detection system**.** The key contribution in this research is the evaluation of the performances of Association rule against Support vector machine based on their accuracy for intrusion detection system using two data sets (NSL-KDD and NUSW-NB15 data set).

2. RESEARCH METHODOLOGY

The UNSW-NB15 data set was developed by using an IXIA Perfect Storm tool in the Cyber Range Laboratory of the Australian Centre for Cyber Security (ACCS) to extract a hybrid of modern normal and modern attack behaviors. This data set involves nine attack categories and 49 features. The attacks type was classified into nine groups which are: Analysis, Dos, Exploit, Fuzzers, Generic, Reconnaissance, Shellcode, Worms, and Backdrop. It contains 82,332 training sets and 175,341 testing set records.

NSL-KDD data set is a refined version of KDD'99, and it contains essential records of the complete KDD'99 data set. NSL-KDD data set contains 4,898,431 numbers of Records. This data set involves four attack categories and 41 features. The attack class presented in NSL-KDD data set are as follows: DOS, Probing, U2R and R2L. figure 1 designates the System architecture for intrusion detection using association rule and support vector machine.

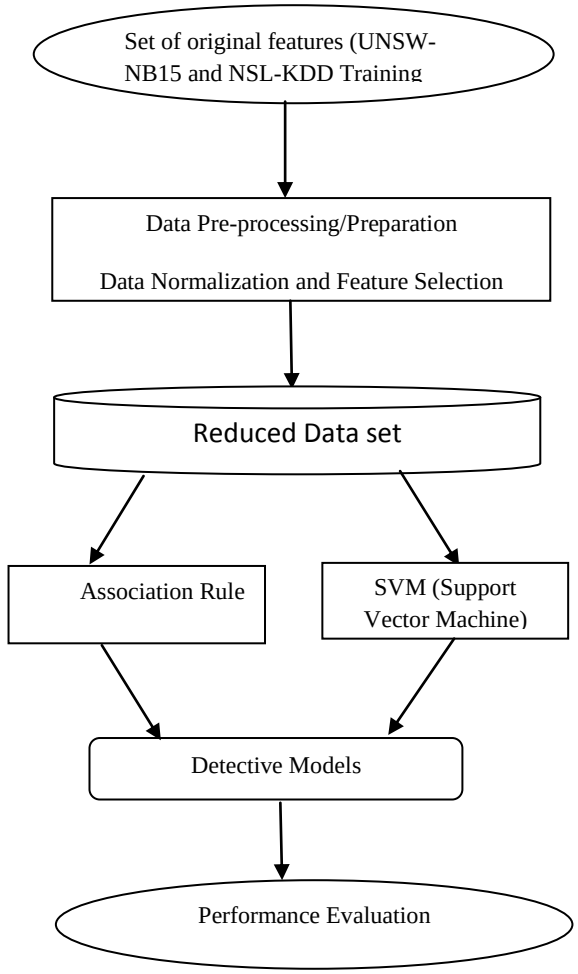

Figure 1: System architecture for intrusion detection using association rule and support vector machine

2.1 Description of feature selection Method

The research will make use of NSL-KDD and UNSW-NB15 data set as stated in equation(1)

$$D = \{S_i, C_j\}; \forall S_i \exists f: f = \{1,2,3,4...,y\} \quad (1)$$

Where D represents the training set containing set of network S_i, i= 1,2, … ,n with an assigned class label C_j, j =1,2,...,m, where n is the number of instances, m is the number of classes and f represents set of features in S_i. These set of features will undergo data pre-processing (since the foundation for successful data mining process is data pre-processing which will be achieved using Feature Scaling (Min-Max) presented in equation (2)

$$X' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2)$$

Where, X is an original data, $X^{'}$ is the normalized data set, X_{min} is the minimum value of X and X_{max} is the maximum value of X.
In order to select the most relevant features from f, the Mutual Information feature selection (CFS) method defined in equation (3) and ANOVA (Analysis of Variance) feature selection techniques will be adopted.

Formally, the mutual information of two discrete random variables X and Y can be defined as:

$$I(X,Y) = \sum_{y\in Y}\sum_{x\in X} P(x,y)\ \log\left(\frac{p(x,y)}{p(x)p(y)}\right) \qquad (3)$$

Where $p(x,y)$ is the joint probability function of X and Y,

And p(x) and p(y) are the marginal probability of X and Y respectively.
Also, the ANOVA is the classical method to compare means of multiple (≥2) groups.
Let x_{ii} be the j_{th} observation from the i_{th} group. Here the number of samples from each group remains the same. Denote *x* as the grand sample mean and x_i as the sample mean of group x_{ij}. Observations can be re-written as:

$$x_{ij} = \bar{x} + (\bar{x_i} - \bar{x}) + (x_{ij} - \bar{x_i}) \qquad (4)$$

This leads to the following model

$$x_{ij} = \mu + \propto_i + \varepsilon_{ij} \qquad (5)$$

Where μ and α_i are grand mean and i_{th} group mean respectively. The error term ε_{ij} is assumed to be *i* differences from a normal distribution.
The null hypothesis in ANOVA implies that all group means are the same as shown below:

$$\propto_1 = \propto_2 = \ldots = \propto_k \qquad (6)$$

The selected features from the two feature selection algorithms will be fed to train Association Rule and Support Vector Machine (SVM) classification algorithm in order to classify normal traffic from attacks. SVM classification function is based on the concept of decision planes that define decision boundaries between classes of samples. SVM performs classification by linearly separating class C ε (c_0, c_1) with a suitable hyperplane using equation (7) and (8)

$$g(x) = w.x + b = 0 \qquad (7)$$

$$|w.x + b| / ||w|| = 1/||w|| \qquad (8)$$

Where *w* represents the weight vector, *b* is the bias factor, and *x* is the data sample. Equation (7) is used to create hyperplanes that separates class C, while equation (8) is used to compute distances (margin) between the classes. The most suitable hyperplane is the one with maximum margin, and this is obtained by minimizing weight vector ||w||. Additionally, an association rule data mining algorithm–Apriori was used to find feature patterns from our data set that exceed the minimum support threshold.

Descriptions

1. Association rule [19]**:** Let I = {i_1, i_2, . . . i_n} be a set of literals call items. Let D be a set of all transactions where each transaction T_i is a set of items such that $T_i \subseteq$ I. Let X, Y be a set of items such that X, Y ⊆ I. Then, an association rule is an implication in the form X ⇒ Y where X, Y ⊂ *I* and $X \cap Y = \emptyset$
2. Support [20]**:** Given that I is a set of m domain items and X = {x_1, x_2, . . ., x_k} be a k-itemset, where X⊆I and 1 ≤ k ≤ m. Let D_p be a precise database of n transactions, each transaction $t_j \subseteq$ I. The support of an itemset X in a transaction t_j, *sup*(*X*, t_j), is equal to 1 if X ⊆t_j or 0 if X t_j.

   The support of an itemset X in an entire database D_p is the frequency with it appears in the database as shown in Eq.9. The *minimum support* is a user defined threshold (δ) that must be reached or exceeded before an item can be considered as frequent or not. Therefore, an itemset X in a database D is a frequent itemset if and only if sup(x) ≥ δ **Support** determines the frequency of row values that denotes the association percentage, as reflected in equation (9), while

$$sup(X) = \sum_{j=1}^{n} \sup(X, t_j) \qquad (9)$$

3. Confidence [21]**:** The confidence of a rule X ⇒ Y is equivalent to the ratio of the support of X ∪ Y to that of X as shown in Eq. 10. It can also be seen as the conditional probability that Y occurs, given that X has occurred.

$$con\ f(X \Rightarrow Y) = \frac{sup(X \cup Y)}{sup(X)} \qquad (10)$$

4. Apriori algorithm [22] uses two steps "join" and "prune" to decrease the search space. It is an iterative approach to determine the most recurrent itemset. It uses a level-wise breadth-first bottom-up technique with a candidate generate-and-test scenario to recognize recurrent patterns from the UNSW-NB15 and NSL-KDD data set. The algorithm first generates candidate patterns of cardinality *k* (i.e., candidate k itemset) and check if each them is recurrent by testing if their support or occurrence meets or exceed the user-defined minimum support threshold (δ). Then, the algorithm produces candidate patterns of cardinality *k=1* (i.e., candidate (k+1)-item sets). This process is performed regularly to determine recurrent patterns of all cardinalities. The implementation was done using python programming and evaluation was based on standard metrics.

2.2 Performance Evaluation for SVM (Support Vector Machine)

The performance metric of the models used for classification was based on the following metrics:

Evaluation Metrics
TP (True Positive): Number of positive instances correctly classified (attack data is detected as an attack)
TN (True Negative): Number of negative instances correctly classified (normal data is detected as an normal)
FP (False Positive): Number of negative instances incorrectly classified as positive (normal data is detected as an attack)
FN (False Negative): The number of positive instances incorrectly classified as negative (attack data is detected as a normal).

$$Accuracy = \frac{TP + TN}{TP + TN + FN} \qquad (11)$$

$$False\ Alarm = \frac{FP}{FP+TN} \qquad (12)$$

2.3 Performance Evaluation of Association Rule

In Association rule, the frequent item sets are searched and found using Apriori algorithm, then mining of association rule is carried out using Association Rule mining (ARM). The Association rules mining algorithm was implemented based on support and confidence being the two basic criteria's in association rule mining.

2.4 Performance Evaluation of SVM vs Association Rule:

Features were selected from UNSW-NB15 and NSL-KDD data set and was trained on Association rule and support vector machine. The performance evaluation of non-rule-based machine learning (SVM) with selected number of features was done, and that of rule based (Association Rule) with selected number of features was also done. The results were compared based on their accuracy.

3. RESULTS AND DISCUSSION

Table 1: Feature Selection Table

Data Set	Feature Selection Method	Number of Feature Selected
NSL-KDD	Mutual information	15
	ANOVA	15
UNSW-NB15	Mutual information	10
	ANOVA	5

3.1 Performance Evaluation of SVM with NSL-KDD

Table 2: Accuracy, precision and False Alarm Rate when 15 features were selected from NSL-KDD data set using Mutual Information and 15 features selected from NSL-KDD data set using ANOVA and trained with Support Vector Machine.

No. of features	Feature selection	Precision	FAR (False Alarm Rate)	Accuracy
15	Mutual Information	0.9058	0.0778	0.7200
15	ANOVA	0.9777	0.0197	0.7965

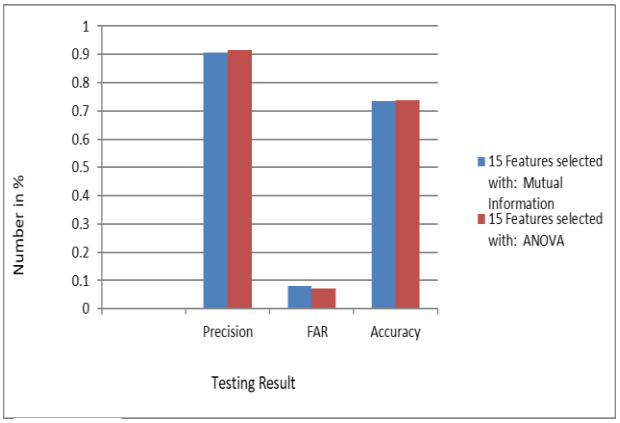

Figure 2: Mutual Information against ANOVA when 15 features were selected from NSL-KDD data set using mutual information feature selection and 15 features selected from NSL-KDD data set using Analysis of Variance feature selection and trained with SVM. With 15 features selected from NSL-KDD data set using mutual information, the accuracy is 72% with 7.7% FAR and the 15 feature selected using ANOVA gives 79% accuracy with 1.9% FAR when trained with SVM

Table 3: Confusion matrix when the features selected from NSL-KDD data set is trained with SVM.

Selection Method	True Negative	False Positive	False Negative	True Positive
NSL-KDD with Mutual Information	8954	756	5554	7278
NSL-KDD with ANOVA	9518	192	4395	8437

Table 4: Accuracy, precision and False Alarm Rate when 10 features were selected using Mutual Information and 5 features selected using ANOVA from UNSW-NB15 Data set and trained with SVM.

No. of features	Feature selection	Precision	FAR (False Alarm Rate)	Accuracy
10	Mutual Information	0.9028	0.2208	0.9041
5	ANOVA	0.9088	0.1853	0.8505

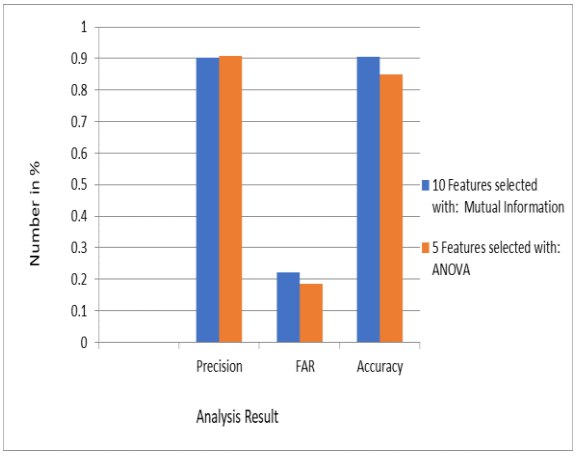

Figure 3: Graph of Mutual Information against ANOVA when 10 features were selected from UNSW-NB15 data set using Mutual Information and 5 features selected from UNSW-NB15 data set using ANOVA.

Table 5: Confusion matrix when the features selected from UNSW-NB15 data set is trained with SVM.

Selection Method	True Negative	False Positive	False Negative	True Positive
UNSW-NB15 with Mutual Information	43632	12368	15373	103968
UNSW-NB15 with ANOVA	45619	10381	15817	103524

3.2 Performance Evaluation of Association Rule with NSL-KDD and UNSW-NB15

Table 6: Support verses Confidence when 15 features were selected from NSL-KDD data set using Mutual Information and 15 features selected from NSL-KDD data set using ANOVA and trained with Associative rule.

No. of Features	Feature Selection	Support	Confidence
15	Mutual Information	0.20	0.66
15	ANOVA	0.20	0.68

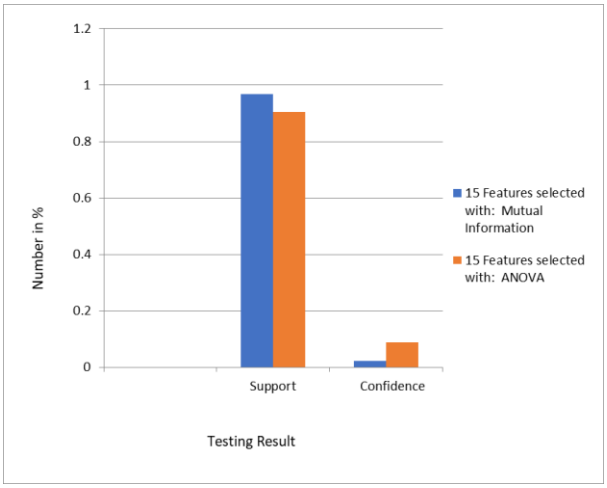

Figure 4: Mutual Information against ANOVA when 15 features were selected from NSL-KDD data set using Mutual Information and 15 features selected from NSL-KDD data set using ANOVA and trained with Associative rule.

Table 7: Confusion matrix when the features selected from UNSW-NB15 data set is trained with Association Rule.

Selection Method	True Negative	False Positive	False Negative	True Positive
UNSW-NB15 with Mutual Information	4709	3982	38189	75297
UNSW-NB15 with ANOVA	7152	159	6390	6459

Table 8: Support versus Confidence when 10 features were selected from UNSW-NB15 data set using Mutual Information and 5 features were selected from UNSW-NB15 data set using ANOVA and trained with Associative rule.

No. of Features	Feature Selection	Support	Confidence
10	Mutual Information	0.05	0.68
5	ANOVA	0.01	0.40

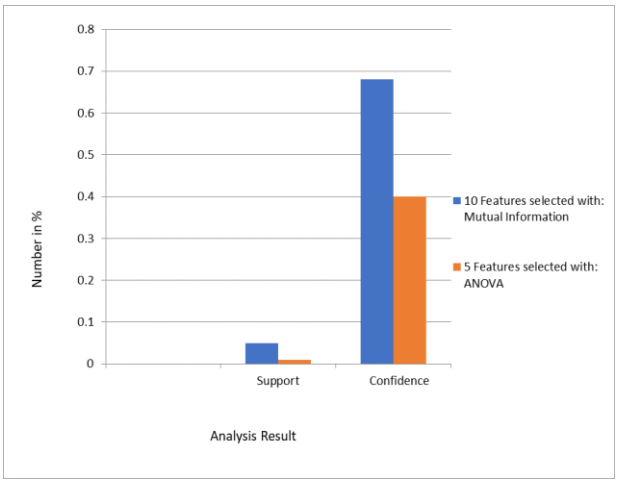

Figure 5: Mutual Information against ANOVA when 10 features were selected from UNSW-NB15 data set using Mutual Information and 5 features selected from NSL-KDD data set using ANOVA and trained with Associative rule.

Table 10: Confusion matrix when the features selected from NSL-KDD data set is trained with Association Rule

Selection Method	True Negative	False Positive	False Negative	True Positive
UNSW-NB15 with Mutual Information	8462	694	7419	6019
UNSW-NB15 with ANOVA	40126	79214	168832	72429

3.3 Performance Evaluation OF SVM against Association Rule

In data mining, accuracy is the skill of the learning algorithm to predict accurately. These defines the percentage of correct predictions made from all predictions. Thus, the accuracy from rule based (Associative rule) is compared against the accuracy from non-rule based (Support Vector Machine). To achieve this, the accuracy of each of the set of features selected from NSL-KDD and UNSW-NB15 data set, when trained with Association Rule and Support Vector Machine was compared against each other as follows:

Table 11: Accuracy of SVM against Accuracy of Association rule with NSL-KDD data set

No. of Features	Feature Selection	Accuracy of SVM	Accuracy of Association rule
15	Mutual Information	0.7200	0.66
15	ANOVA	0.7968	0.68

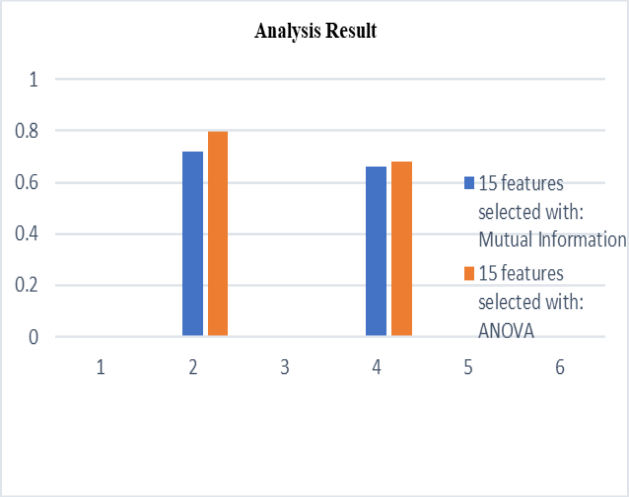

Figure 6: Graph of SVM against Association Rule

From **figure 6** above, the accuracy of rule-based (Association rule) is 66% with mutual information and 68% with ANOVA while the accuracy of SVM is 72% with mutual information and 79% with ANOVA. Thus, SVM (non-rule-based machine learning) with both filter and wrapper method perform excellently in terms of accuracy than Association rule which is a rule-based machine learning.

3.4 Performance Evaluation of SVM against Association Rule.

Table 12: Accuracy of SVM against Accuracy of Association Rule when features were selected using NUSW-NB15 Data set

No. of Features	Feature Selection	Accuracy of SVM	Accuracy of Association rule
10	Mutual Information.	0.9041	0.68
5	ANOVA	0.8505	0.40

3.5 Performance Evaluation of SVM against Association Rule.

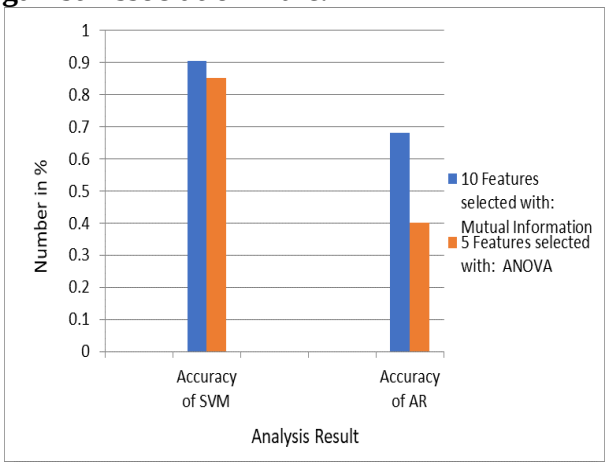

Figure 7: Accuracy of SVM against Accuracy of Association with NSL-KDD data set

From figure 7 above, the accuracy of rule-based (Association rule) is 68% with mutual information and 40% with ANOVA, while the accuracy of SVM is 90% with mutual information and 85% with ANOVA. Thus, non-rule-based machine learning is of higher accuracy than the rule based.

3.6 Contribution to Knowledge

This research work evaluates the accuracy of Association rule and Support vector machine using the two data set (KDD '99 and UNSW data set).

4. CONCLUSION AND RECOMMENDATION

In this research, the result of a rule based (Association Rule) was compares against a non-rule based (Support Vector Machine) machine learning algorithm, using filter method(mutual information) and Wrapper method (Analysis of Variance) as the feature selection method, in which the features were selected from NSL-KDD data set and UNSW-NB15 data set and trained using Association Rule and support Vector Machine.

From the results Mutual information with NSL-KDD selects 15 features and trained with SVM which gives 72% accuracy, while ANOVA with NSL-KDD also selects 15 features and trained with SVM which give 79% accuracy.

Also, Mutual information with NUSW-NB15 selects 10 features and trained with SVM which gives 90% accuracy, while ANOVA with NUSW-NB15 selects 5 features and trained with SVM, and it gives 85% accuracy.

In addition, Mutual information with NSL-KDD selects 15 features and trained with Association Rule which gives 66% accuracy, while ANOVA with NSL-KDD also selects 15 features and trained with Associative rule which gives 68% accuracy.

Also, Mutual information with UNSW data set selects 10 features and trained with Association Rule which gives 68% accuracy, while ANOVA with UNSW selects 5 features and trained with Associative Rule and it gives 40% accuracy.

From the above results, SVM (non-rule-based machine learning) with both filter and wrapper method perform excellently in terms of accuracy than Association rule a rule-based machine learning. Thus, the non-rule-based machine learning perform excellently than the rule base.

5. REFERENCES

[1] J. Jang-Jaccard, S. Nepal 2014 A survey of emerging threats in cybersecurity, J. Comput. Syst. Sci. 80 (5) (2014) 973–993, doi: 10.1016/j.jcss.2014.02.005

[2] Colesky, M., J.-H. Hoepman, and C. Hillen, 2016 "A critical analysis of privacy design strategies", IEEE Security and Privacy Workshops (SPW), first online O4 August 2016, doi:10.1109/SPW.2016.23

[3] Cybercrime 2012 Cyber Crime & Security survey. Commonwealth of Australia, http://realbusiness.co.uk.

[4] Hung-Jen Liao, Chun-Hung Richard Lin, Ying-Chih Lin, Kuang-Yuan Tung. 2013. Intrusion detection system: A comprehensive review. Journals of network and computer Applications, Vol. 36, pp. 16-24.

[5] Peddabachigari, S., Abraham, A., Grosan, C., and Thomas, J. 2007 Modelling Intrusion Detection System

Using Hybrid Systems. J. Network Computer. Application, 30(2), 114–132.

[6] Aladesote O. I., Boniface K. A., and Folasade D. 2014 Intrusion Detection Technique using Hypothesis Testing. Proceedings of the World Congress on Engineering and Computer Science, 1(2), 978-988.

[7] Bowker, Art 2012 The Cybercrime Handbook for Community Corrections: Managing Risk in the 21st Century. Springfield: Thomas. ISBN 9780398087289. Archived from the original on 2 April 2015. Retrieved 25 January 2015.

[8] O. Pathan, A. Sakib 2014 The state of the art in intrusion prevention and detection, pp. 335–360.

[9] Bace, R. and Mell, P. 2001 Intrusion Detection Systems, National Institute of Standards and Technology (NIST), TechnicalReport,bhttp://www.nist.gov/manuscript.publication.search.cfm? pub.

[10] O.O. Olasehinde, B.K. Alese, A.O. Adetunmbi, 2018 Performance evaluation of bayesian classifier on filter-based feature selection techniques, Int. J. Comput. Sci. Telecommun. 9 (7) 24–30.

[11] Tsai, F. S. and Chan, K. L. 2009 Blog Data Mining for Cyber Security Threats, in: Data Mining for Business Applications, 169–182.

[12] Marano N.S., Betanzos A.A., Estevez R.M. 2009, A Wrapper Method for Feature Selection in Multiple Classes Data sets. IWANN '09 Proceedings of the 10th International Work-Conference on Artificial Neural Networks: Part I: Bio-Inspired Systems: Computational and Ambient Intelligence. Pp 456-463.

[13] Fatogun, B. A. 2012 Denial of Service Attack Detection Using Machine Learning Techniques, A Thesis in the Department of Computer Science, Federal University of Technology, Akure, Nigeria, pp 2-4.

[14] Devikrishna K S, and Ramakrishna B. B. 2014 An analysis of Intrusion Detection System using Back Propagation Neural Network. International Journal of Engineering Research and Applications (IJERA), 3(4), 1959-1964.

[15] Liu, H. Y., Xiangdong C., and Shalini L. 2017, Understanding Modern Intrusion Detection Systems: A Survey. College of Technology, Eastern Michigan University, United States 2(1), 1-9

[16] Nour, M. and Jill, S. 2015 A Hybrid Feature Selection for Network Intrusion Detection Systems: Central points. The Proceedings of the 16th Australian Information Warfare Conference, pp. 5-13.

[17] S.Q. Qais, J.A. Mohd, M.Z. Abdullah 2016 Anomalies classification approach for network-based intrusion detection system, Int. J. Netw. Secur. 18 (6) 1159–1172.

[18] Youn, E. and Jeong, M. K. 2009 Class Dependent Feature Scaling Method using Naive Bayes Classifier for text data mining. Pattern Recognition Letters. 30 (5), 477–485.

[19] R. Agrawal, T. Imielinski, A. Swami 2013 Mining association rules between sets of items in large databases, In proceedings of ACM SIGMOD conference 207–216.

[20] A. Cuzzocrea, C.K. Leung, R.K. MacKinnon 2015 Approximation to expected support of frequent itemset in mining probabilistic sets of uncertain data, Procedia Comput. Sci. 613–622.

[21] K. Lai, N. Cerpa, Support vs Confidence in Association Rule Algorithms, in: Conference of the Chilean Operations Research Society, 2001, pp. 1–15.

[22] C.K. Leung 2015 Big data mining applications and services, In Big data application and services BigDAS 1–8.