ISSN 2278-3091
Volume 9, No.5, September - October 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse60952020.pdf
https://doi.org/10.30534/ijatcse/2020/60952020

# DDoS Attack Classification on Cloud Environment Using Machine Learning Techniques with Different Feature Selection Methods

**C.Bagyalakshmi[1] , Dr.E.S.Samundeeswari[2]**
[1]Research Scholar, Department of Computer Science, Vellalar College for Women, Erode, Tamilnadu, bagyachithra@gmail.com
[2]Associate Professor, Department of Computer Science, Vellalar College for Women, Erode, Tamilnadu.

## ABSTRACT

Cloud Computing is a prominent compelling paradigm for managing and delivering services over the Internet. It is modifying the landscape of information technology in terms of data storage. In large data storage requirements, highest priority is to be given for data security. Intrusion is one of the important security issues in today's cyber world. Due to networked nature of the cloud, resources, data and applications are vulnerable to the attack in cloud environment. Intrusion Detection Systems (IDS) are employed in the cloud to detect malicious behavior in the network and in the host. Distributed Denial of Service (DDoS) attack is one of challenging task in IDS, as it creates a huge volume of malicious data in the network. Data mining methods for cyber analytics provide support for intrusion detection. A significant number of techniques are developed, based on machine learning approaches. Feature selection methods also play an important role in reducing the dimensionality of the dataset. In this work, two approaches are proposed and the dataset is collected from NSL-KDD. The first approach uses Learning Vector Quantization (LVQ), a filter method and the second approach uses Principal Component Analysis (PCA), a dimensionality reduction method. The selected features from each approach is used for classification using Naïve Bayes (NB), Support Vector Machine (SVM) and Decision Tree (DT) and compared the results in terms of their detection capability for DDoS attack. Results shows that LVQ based DT technique overtakes the others in terms of attack identification.

**Key words:** Cloud Computing, SVM, DT, NB, LVQ, PCA.

## 1. INTRODUCTION

Cloud computing is an emerging technical advancement for providing information technology in terms of Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). Adoption of cloud computing is considered as public, private and hybrid clouds [1]. Public clouds basically reside outside of an organization's premises and is accessible through the Internet. Private clouds reside in an organization's premises. Hybrid clouds allow organizations to use a mix of private and public clouds to provide services. For organizations in order to change their working environment to clouds, it becomes important for cloud providers to assure major level of security for their clients. Cloud service providers are in need to ensure the security measures through firewalls and IDS by enhancing the architecture of cloud to the clients [2][20].

Intrusion reduces the confidentiality, availability and integrity of entire computer resources. An attackers causes threats to network security by means of trespassing unauthorized users, stealing of assets, acquisition of privileges, performing beyond the limit and injecting malicious activity into the network traffic. In order to solve the above mentioned problems, IDS have been developed to analyze/monitor the network traffic and test whether the traffic is normal or variant [7]. Due to variations in network configurations several types of intrusion detection types are emerged. Each type of IDS has different advantages and disadvantages in terms of cost, configuration and detection rate. Based on the deployed area, IDS is classified as host based IDS, network based IDS, hypervisor based IDS and distributed IDS.

Data mining tools are used to fetch the details about the different type of attacks [3][21]. Inputs from the mining tools help to improve the data security and to identify the behavior of the attackers. Behaviors of the attack is detected through supervised learning and unsupervised learning. Data mining approach helps to improve the network security by collecting the information on the activities of the attackers. Based

on obtained data appropriate algorithms applied to predict behavior of attacker in the network.

In this research work, machine learning techniques are implemented on NSL-KDD dataset to detect DDoS attack in cloud environment. For implementing the mining methods effectively, LVQ and PCA based feature selection techniques are used. Classification techniques are implemented on selected features.

## 2.BACKGROUND STUDY

Different data mining techniques are to be studied in IDS. Literatures are collected by focusing towards the identification of different classification techniques to implement IDS for cloud environment. Background study about these topics are discussed in this section.

Jaswinder et.al.,[6](2012), focused on Distributed Denial of Service (DDoS) attack and identified different types of DDoS attacks by simulation. Legitimate traces are found during simulation of DDoS attacks. Network topology is simulated along with attached real time traces. The impact of attack is measured in terms of metrics like throughput and percentage link utilization.

Singh et.al., [3](2013), developed a predictive model for supervised learning based Intrusion Detection System (IDS) to identify the intruders and attackers in a network. The Naïve Bayes (NB) and J48(C4.5) techniques are implemented on a standard benchmark dataset. The NB algorithm gives better accuracy while compared with tree based supervised learner J48. The supervised learning based IDS provides a proper way to identify the intruders and hackers in real time networks.

Ryan et.al., [12] (2013), created a set of experiments to examine the performance of most typical virtualization techniques under typical Denial of Service (DoS) attacks. However, a small packet is sent at a high rate caused degradation to a virtualized system. DoS attack on a virtualized system has serious performance impacts. Hence, isolated environment suffered from greater performance degradation compared with its non -virtualized counter environment.

Keisuke et.al., [8] (2014), developed an intelligent detection system for Distributed Denial of Service (DDoS) attacks. It detects attack patterns using network packet analysis and utilizing machine learning techniques. Dataset provided by the Center for Applied Internet Data Analysis is utilized with detection system using a Support Vector Machine in Radial Basis Function (Gaussian) kernel. Calculated bytes per second with the time sequence in DDoS attack and detection system is accurate in detected DDoS attacks.

Zerina et.al., [17] (2017), proposed an automated classification system for Denial of Service (DoS) attack detection in the cloud computing. This study is performed in several phases like attack simulation, data collection, feature selection, and classification. Data for this study is obtained by simulating the cloud environment and DoS attack, using Wireshark with Tshark option is extracted with necessary features. Support Vector Machine (SVM) is one of the model for classification of DoS attacks and normal network behaviors.

## 3.METHODOLOGY

This section describes the detection of DDoS attack in NSL-KDD dataset using data analysis. The work flow of intrusion detection model is shown in Figure 1. Two feature selection methods are applied to obtain a reduced set of features.

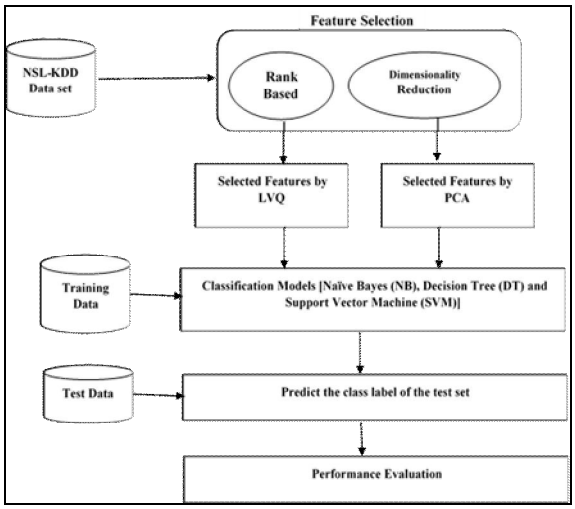


**Figure 1:** Proposed Model for IDS in Cloud Environment

The reduced feature set is applied to various classification techniques. Validation metrics are calculated from the confusion matrix.

### 3.1 Data set Description

The cost of network erection in a real distributed testing environment is very high. Simulation is a significant method in network research as it can be used to analyze network related problems under different protocols, traffic and topologies with less cost. There are two types of data sets available, such as direct data set and public data set. The dataset created with the help of some open source software is called direct data set, if the user donated their own

dataset with help of online platform then it is called as public dataset.

A public dataset, NSL-KDD is used in this work [18]. The original data set contains 2, 26,283 instances with 42 features and 4 category of attacks. For this analysis work, only DDoS type of attacks are considered and it results in 15,452 instances of DDoS attack. Data set is divided into 70% for training and 30% for testing. R tool is used to implement the above model.

### 3.2 Feature Selection

Feature selection method is a technique which selects the important features that has more impact on the predicted variable. Inclusion and exclusion of data in feature selection will not affect the entire data set for prediction. In the proposed work two methods of feature selection are carried out. They are filter method and dimensionality reduction method.

#### 3.2.1 Filter Method

LVQ algorithm is supervised learning and it is used in Artificial Neural Network. The architecture of LVQ is n number of input units and m number of output units. The layers are fully interconnected with having weights on them. The working principle of LVQ method is based on k-Nearest Neighbor (k-NN) algorithm [5]. The LVQ parameters used for training process are x,T,wj,Cj and j. x is a training vector x(x1,x2……xn). T is the class for training vector x. wj is weight vector for jth output unit. Cj is the class associated with the jth output unit.

The following steps are involved in LVQ algorithm,

**Step 1:** Initialize, determine the initial weight, maximum epoch (number of training processes to be repeated) and the learning rate (alpha) value.

**Step 2:** If repetition conditions are fulfilled, do steps 2- 8.

**Step 3:** Set initial conditions epoch= 0.

**Step 4:** If the condition (epoch<MaxEpoch) then epoch= epoch+ 1

**Step 5:** Calculate the minimum distance ||xi-wj|| using Euclidean distance.

**Step 6:** Update weight wj with the conditions:

If T=Cj, then wj (new) = wj(old) + α(x – wj(old)) (closer together)

If T≠Cj, then wj (new) = wj(old) –α(x– wj(old)) (further apart)

**Step 7:** Reduce learning rate (α) = $\alpha - (0{,}1 * \alpha)$

**Step 8:** Stop condition test: the condition where the learning rate (α) and the error reach the specified target value.

#### 3.2.2 Dimensionality reduction method

Dimensionality reduction is a method for reducing irrelevant features in sequential manner. Dimensionality reduction is implemented using Principal Component Analysis (PCA). Large set of data variables are reduced as small variables using PCA [19]. The data set patterns are easily highlighted for visualizing the data variables. PCA uses variable distribution in an orthogonal statistical format in order to carry out transformation of data variables. PCA concepts are implemented based on mathematical Standardization, Covariance, Eigenvalues and Eigen vectors.

**Step 1: Standardization**

Standardization is one of the method of mathematical structure followed to implement PCA algorithm in larger data set. Data initially scaling is to be carried out between 0 to1, all the variables are equally contributing for the analysis. Mathematically standardization is carried out using equation (1), after standardization all variables are converted into same scale.

$$S = \frac{x-\mu}{\sigma} \qquad (1)$$

Where,

x – Number of observation in the dataset,

μ - Mean of total observed data and

σ – Standard deviation of observed data

**Step 2: Covariance**

Covariance is calculated by considering the variables of input data set, its variations are relevant to mean. Highly correlated variables expresses the irrelevant data. A covariant symmetrical type of matrix is developed based on the number of dimensions. This matrix gives all possible combinational pairs are represented with respect to covariance. The relation between two variables are represented by following conditions, if two variables are directly correlated then positive otherwise if two variables are inversely correlated then negative(2)

$$\sigma^2_{xy} = \frac{\sum_i (x_i - m_x)(y_i - m_y)}{n-1} \qquad (2)$$

Where,

$x_i$ & $y_i$ – Input values (attributes) from the data set

$m_x$ & $m_y$ – Mean values (attributes) of the data set

**Step 3 – Eigen Values and Eigen Vectors**

Eigen values and Eigen vectors play an important role in the operation of PCA. Eigen values are diagonal numbers of a covariance matrix whereas Eigen vectors are new rotated axes of covariance matrix. In order to calculate Eigen values, initially the covariance matrix is computed using equation (3),

$$|A - \lambda \cdot I| = 0 \qquad (3)$$

Where,

A – Input value of covariance matrix, λ – Eigen value and I – Identity Matrix

A new variable has been created for the PCA component with some information loss. Dimensionality reduction is carried out using PCA

with equal number of original values. Based on the ranking order of Eigen vectors and their values from highest to lowest order of significance in PCA. After completing the above set of procedures to find the Eigen values and Eigen vectors, some more additional steps are to be followed to complete the PCA implementation in the data set. The additional steps includes transposition of Eigen vectors and transpose the adjusted data. By these adjusted data, new data can be formulated using equation (4),

$$y = A^T \cdot x \quad (4)$$

Where, AT – Transpose of Eigen value input vectors and x – Input data set.

### 3.3 Classification Techniques

Data analysis is carried out using classification algorithms in NSL-KDD data set. Classification algorithms will support prediction in the data set. In classification techniques, a model is developed in connection with the relationship between values of predictors and target [10]. Based on the relation, a model can be used for different set of data for which classes are unknown. In this work, classifiers used are Naïve Bayes, Support Vector Machine and Decision Tree. These classification models are used to predict the normal and malicious record in data set.

#### 3.3.1 Naïve Bayes (NB)

Naive Bayes is constructed on theory of probability machine learning algorithm based on the Bayes Theorem. It is used in wide variety of classification tasks. The probabilities are calculated for each feature and highest probability is selected. Naive Bayes classifier assumes that all features are not related to each other. The probability of malicious records for the dataset are calculated as shown in the equation (5).

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (5)$$

Where,

- P (B|A) - Probability of the evidence (Likelihood).
- P (A|B) - Probability of the hypothesis gives that the evidence is there (Posterior Probability)
- P (A) - Number of probabilities of attack in the data set (Class Prior Probability)
- P (B) - Total number of available features (Predictor Prior Probability)

On implementing NB algorithm for the dataset with LVQ and PCA based feature selection gives and accuracy of 0.9197 and 0.8721 respectively.

#### 3.3.2 Support Vector Machine (SVM)

In a set of large data with different class members a decision plane is required to separate. SVM technique supports decision plane for defined decision boundaries. It takes the data as an input and output as a line that separates those classes if possible. To find the data points closest to the line from both classes, these data points are called support vectors. The hyperplane and support vectors distance are computed, is called as a margin. The hyperplane for which the margin in maximum is the optimal hyperplane [18]. For the given data set SVM method was experimented and it gives an accuracy values of 0.9288 in LVQ method and 0.9847 in PCA method.

#### 3.3.3 Decision Tree (DT)

Decision Tree (DT) is a classification and prediction based method. Decision tree algorithm consists of predicted decisions with conditional control statements. All possible consequences of decisions are considered while developing this algorithm. DT includes several nodes and subsets of all possible decisions and outcomes [3]. DT algorithm is a better method of representing the data it considers all probable part of final decision through tree-like structure. While carrying out this process recursive mode decision making is conducted. It is also able to handle high dimensional data with good accuracy. DT algorithm results on the given data set reveals that an accuracy value of 0.9874 in LVQ method and 0.9860 in PCA method.

## 4.RESULT AND DISCUSSION

### 4.1 Summary of Results

Results are obtained R tool is used for predicting the malicious record from the NSL-KDD data set. Feature selection methods are applied, the resultant feature set is used for classification. The validation metrics are calculated using the standard formulas of accuracy, precision, recall, specificity and f-measure, and tabulated in the Table 2 & 4.

Precision – proportion of normal record (positive cases) that were correctly identified.

Recall or Sensitivity – proportion of actual normal record (positive cases) which were correctly identified

Specificity – proportion of actual malicious record (negative cases) were correctly identified

F-Measure - F-Measure is combination of results obtained from Recall and Specificity

### 4.2 LVQ Based Classification

After implementation of LVQ method in the data set 20 important variables (Table 1) as shown in figure 2 are filtered out from 41 variables of the given data set.

**Table 1:** Implementation of LVQ ranking for most important variables

| S.No | NSL-KDD Attribute Names | LVQ Ranking (out of 41) | S.No | NSL-KDD Attribute Names | LVQ Ranking (out of 41) |
|---|---|---|---|---|---|
| 1 | same_srv_rate | 0.83733 | 11 | dst_host_srv_serror_rate | 0.22677 |
| 2 | dst_host_same_srv_rate | 0.68643 | 12 | serror_rate | 0.2257 |
| 3 | dst_host_srv_count | 0.61586 | 13 | dst_host_serror_rate | 0.22522 |
| 4 | dst_bytes | 0.57396 | 14 | srv_serror_rate | 0.22449 |
| 5 | dst_host_srv_rerror_rate | 0.51681 | 15 | dst_host_count | 0.21432 |
| 6 | dst_host_rerror_rate | 0.51325 | 16 | service | 0.09433 |
| 7 | rerror_rate | 0.51265 | 17 | srv_diff_host_rate | 0.07459 |
| 8 | logged_in | 0.51233 | 18 | diff_srv_rate | 0.05513 |
| 9 | srv_rerror_rate | 0.51196 | 19 | dst_host_diff_srv_rate | 0.03789 |
| 10 | count | 0.33889 | 20 | dst_host_srv_diff_host_rate | 0.03271 |

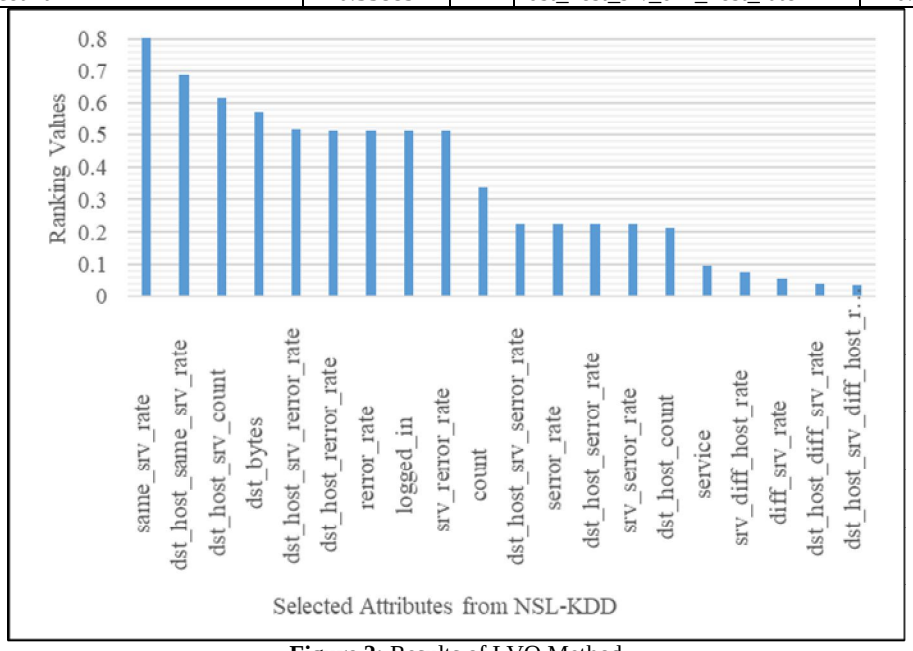


**Figure 2:** Results of LVQ Method

### 4.2.1 Classification on features obtained through LVQ method

Experimental results are shown in Table 2 and figure 3 for the given data set. After implementation of LVQ, Classification techniques are applied such as NB, SVM and DT. The DT classifier provides better performance level as compared to NB and SVM for detecting the malicious record.

**Table 2:** Results of LVQ Method

| | Naïve Bayes | SVM | Decision Tree |
|---|---|---|---|
| Accuracy | 0.9197 | 0.9288 | **0.9874** |
| Precision | 0.9085 | 0.9060 | 0.9887 |
| Recall | 0.9727 | 0.9897 | 0.9914 |
| Specificity | 0.8250 | 0.8249 | 0.9808 |
| F-Measure | 0.9395 | 0.9460 | 0.9901 |

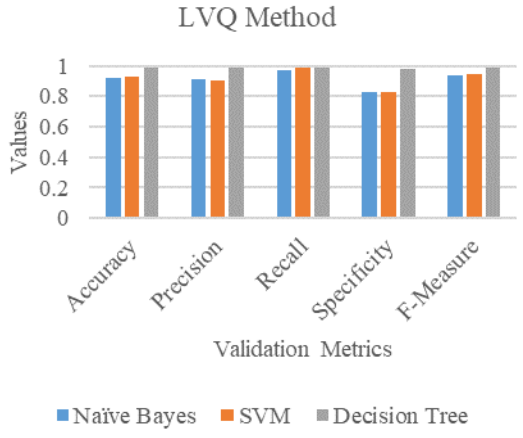


**Figure 3 :** LVQ Method

### 4.3 PCA Based Classification

The importance of variables are filtered out from the data set (Table 3). PCA reduces 21 out of 41 variables in the given data set. It is shown in figure 4.

**Table 3:** List of important variables retrieved by PCA

| S.No | Selected Attribute Names | Dimension Values |
|---|---|---|
| 1. | srv_serror_rate | 18.46 |
| 2. | serror_rate | 18.05 |
| 3. | dst_host_srv_serror_rate | 17.98 |
| 4. | dst_host_srv_rerror_rate | 17.43 |
| 5. | dst_host_serror_rate | 17.39 |
| 6. | dst_host_rerror_rate | 17.18 |
| 7. | srv_rerror_rate | 12.23 |
| 8. | rerror_rate | 11.68 |
| 9. | same_srv_rate | 10.53 |
| 10. | num_root | 8.52 |
| 11. | num_compromised | 8.45 |
| 12. | hot | 8.42 |
| 13. | dst_host_same_srv_rate | 8.38 |
| 14. | su_attempted | 7.52 |
| 15. | logged_in | 5.95 |
| 16. | dst_host_srv_count | 5.89 |
| 17. | srv_count | 5.36 |
| 18. | protocol_type | 4.95 |
| 19. | Count | 4.64 |
| 20. | num_file_creations | 3.78 |
| 21. | dst_host_same_src_port_rate | 3.01 |

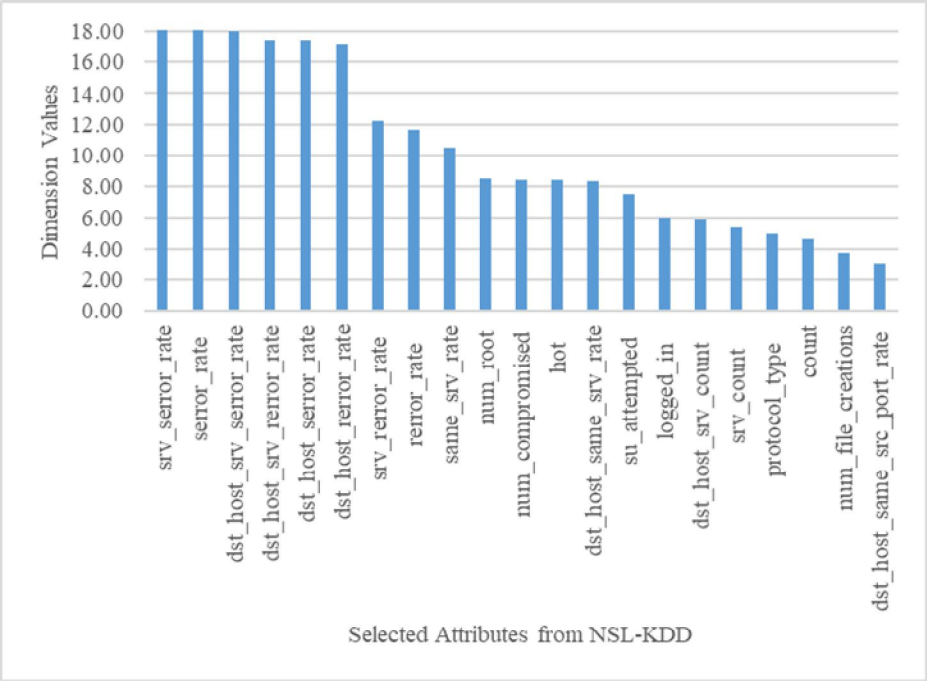


**Figure 4:** Results of PCA Method

### 4.3.1 Classification on features obtained through PCA method

Dimensionality Reduction is implemented using PCA method, the results are shown in figure 5. From the Table 4 performance of DT algorithm with 0.9860 of detection accuracy is better, compared to NB and SVM. In this feature selection method, number of features are considered to be 22 out of 42.

**Table 4:** Results of PCA Method

| | Naïve Bayes | SVM | Decision Tree |
|---|---|---|---|
| Accuracy | 0.8721 | 0.9847 | **0.9860** |
| Precision | 0.9562 | 0.9904 | 0.9983 |
| Recall | 0.8561 | 0.9851 | 0.9781 |
| Specificity | 0.9358 | 0.9839 | 0.9972 |
| F-Measure | 0.9034 | 0.9878 | 0.9881 |

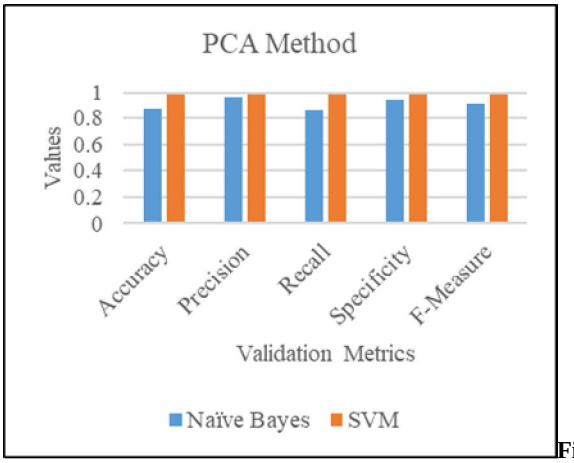


**Figure 5:** PCA Method

### 4.4 Comparative Result

The performance of classifiers were compared, based on the results (Table 5), DT shows (Figure 6) better performance in both feature selection methods. Therefore, LVQ based DT algorithm is preferred to classify the malicious records.

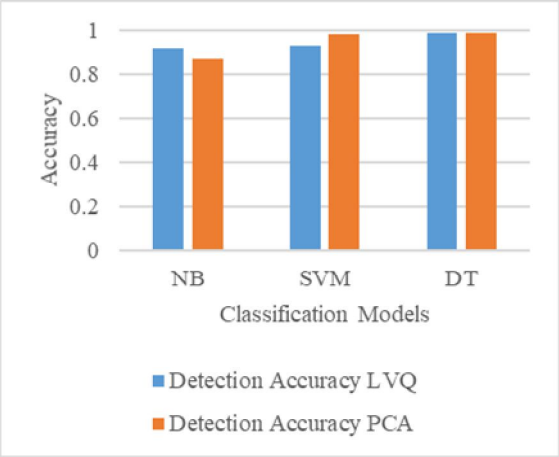


**Figure 6:** Comparative Results of LVQ and PCA

**Table 5:** Comparative Results of LVQ and PCA

| Classification Algorithms | Detection Accuracy | |
|---|---|---|
| | **LVQ** | **PCA** |
| **NB** | 0.9197 | 0.8721 |
| **SVM** | 0.9288 | 0.9847 |
| **DT** | **0.9874** | 0.9860 |

## 5.CONCLUSION

Intrusion detection is a primary part of the data security process. Intrusion detection system for cloud environment data is implemented on a benchmark data set NSL-KDD. Only records related to DDoS attack is considered in this work. The attacks were classified using machine learning techniques of NB, SVM and DT with feature selection methods such as LVQ and PCA. The performance of these algorithms were analyzed to classify the DDoS attack. 20 features out of 42 features were selected by LVQ and 21 features were selected by PCA. Results shows that LVQ based feature selection in DT model identifies the attacks more accurately then the other methods considered. Also works out to be better than other models in terms of precision, recall, specificity and f-score.

## REFERENCES

1. Ahmed Shawish and Maria Salama,"**Cloud Computing: Paradigms and Technologies**", Inter-Cooperative Collective Intelligence: Techniques and Applications, Studies in Computational Intelligence, Springer, DOI: 10.1007/978-3-642-35016-0_2, P.No:39-67, 2014
2. Anna L. Buczak, and Erhan Guven,"**A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection**", IEEE Communications Surveys & Tutorials, Vol:18, Issue: 02,P.No : 1153-1176,2016.
3. Asir Antony Gnana Singh. D and Jebamalar Leavline . E, "**Data Mining In Network Security - Techniques & Tools: A Research Perspective**", Journal of Theoretical and Applied Information Technology, Vol: 57, No.2, ISSN: 1992-8645 & E-ISSN: 1817-3195, P.No: 269-278, 2013.
4. Carlos E. Pedreira**," Learning Vector Quantization with Training Data Selection**", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol: 28, Issue: 01, P.No:157-162, 2006.
5. Ikram.S.T. and Cherukuri.A.K, "**Improving accuracy of intrusion detection model using PCA and optimized SVM**", Journal of Computing and Information Technology, Vol. 24, No. 2, P.No : 133–148, DOI: 10.20532/cit.2016.1002701.2016
6. Jaswinder Singh, Krishan Kumar, Monika Sachdeva and Navjot Sidhu " **DDoS Attack's Simulation using Legitimate and Attack Real Data Sets**", International Journal of Scientific & Engineering Research, Vol:3, Issue: 6, ISSN:2229-5518, P.No:1-5,2012
7. Jian-Hui Wu1, Wei Wei ,Lu Zhang, Jie Wang, Robertas Dama.Evi.Ius , Jing Li, Hai-Dong Wang, Guo-Li Wang, Xin Zhang, Ju-Xiang Yuan and Marcin Wozniak, "**Risk Assessment of Hypertension in Steel Workers Based on LVQ and Fisher-SVM Deep Excavation**", IEEE Access, Special Section on New Trends in Brain Signal Processing and Analysis,Vol : 7, P.No :23109-23119, DOI: 10.1109/Access.2019.2899625,2019.
8. Keisuke Kato and Vitaly Klyuev,"**An Intelligent DDoS Attack Detection System Using Packet Analysis and Support Vector Machine**", International Journal of Intelligent Computing Research (IJICR), Vol:5, Issue:3, ISSN : 2042-4655, P.No: 464-471, 2014.
9. Preeti Mishra, Vijay Varadharajan , Uday Tupakula and Emmanuel S. Pilli ,"**A Detailed Investigation And Analysis of Using Machine Learning Techniques for Intrusion Detection**", IEEE Communications Surveys & Tutorials , 1553-877x, DOI:10.1109/Comst.2018.2847722,2018.
10. Qiao Tian, Jingmei Li and Haibo Liu "**A Method for Guaranteeing Wireless Communication**

**Based on a Combination of Deep and Shallow Learning**", IEEE Access : Special Section On Artificial Intelligence For Physical-Layer Wireless Communications, Vol : 7, P.No :38688-38695, DOI:10.1109/ACCESS.2019.2905754,2019

11. Ruggero Donida Labati, Angelo Genovese, Vincenzo Piuri, Fabio Scotti and Sarvesh Vishwakarma," **Computational Intelligence in Cloud Computing**", P.No : 111-117, doi:10.1007/978-3-030-14350-36.
12. Ryan Shea and Jiangchuan Liu" **Performance of Virtual Machines Under Networked Denial of Service Attacks: Experiments and Analysis**", IEEE Systems Journal, Vol:7, No:2, DOI: 10.1109/JSYST.2012.2221998, P.No:335-345, 2 0 1 3.
13. Shibin David, "**Efficient intrusion detection using machine learning techniques**", Journal of Advanced Research in Dynamical & Control Systems, Vol. 10, Special Issue: 03, ISSN 1943-023X, 2018.
14. Uttam Kumar and Bhavesh N. Gohil,"**A Survey on Intrusion Detection Systems for Cloud Computing Environment**", International Journal of Computer Applications (0975 – 8887), Vol: 109, Issue: 01, P.No : 6-15, 2015.
15. Virgil D. Gligor,"**A Note on Denial-Of-Service in Operating Systems**", IEEE Transactions on Software Engineering, Vol: 10, Issue: 03, P.No:320-324, May 1984.
16. Yang Degang, Chen Guo, Wang Hui And Liao Xiaofeng,"**Learning Vector Quantization Neural Network Method for Network Intrusion Detection"**, Wuhan University Journal of Natural Sciences,Vol.12, No.1, P,No :147-150, Doi 10.1007/S11859-006-0258-Z,2007.
17. Zerina Mašetić, Dino Kečo, Nejdet Doğru and Kemal Hajdarević**," SYN Flood Attack Detection in Cloud Computing using Support Vector Machine**", TEM Journal. Volume 6, Issue 4, ISSN 2217-8309, P.No: 752-759, 2017.
18. Dataset (https://www.unb.ca/cic/datasets/nsl.html)
19. Jason Brownlee,"**Learning Vector Quantization for Machine Learning Algorithms**", 2016 (https://machinelearningmastery.com/learning-vector-quantization-for-machine-learning/)
20. Apoorva Deshpande and Ramnaresh Sharma," **Multilevel Ensemble Classifier using Normalized Feature based Intrusion Detection System**", International Journal of Advanced Trends in Computer Science and Engineering,Volume 7, Issue 5, ISSN 2278-3091,P.No: 72-76,2018
21. Hesham Abusaimeh,"**Security Attacks in Cloud Computing and Corresponding Defending Mechanisms**", International Journal of Advanced Trends in Computer Science and Engineering, Volume 9, Issue 3, ISSN 2278-3091,P.No: 4141-4148,2020