**Intrusion Detection System Using PCA withRandom Forest Approach**

**ABSTRACT:**

With the evolution in wireless communication, there are many security threats over the internet. The intrusion detection system (IDS) helps to find the attacks on the system and the intruders are detected. Previously various machine learning (ML)techniques are applied on the IDS and tried to improvethe results on the detection of intruders and to increasethe accuracy of the IDS. This paper has proposed anapproach to develop efficient IDS by using the principal component analysis (PCA) and the random forest classification algorithm. Where the PCA will help to organize the dataset by reducing the dimensionality of the dataset and the random forest will help in classification. Results obtained states that the proposed approach works more efficiently in terms of accuracy as compared to other techniques like SVM, Naive Bayes, and Decision Tree. The results obtained by proposed method are having the values for performance time (min) is 3.24 minutes, Accuracy rate (%) is 96.78 %, and the Error rate (%) is 0.21 %.

**EXISTING SYSTEM:**

* Iftikhar Ahmad et. al, studied various machine learning algorithms for the intrusion detection system. They compared some of the techniques like SVM, Extremelearning machine and the random forest. The author shave stated the results as the Extreme machine learning method performs a way better as comparedto other algorithms.
* B. Riyazet. al., here worked to improve the quality ofthe dataset to provide it to the intrusion detectionsystem. They have used a fuzzy rule-based featureselection technique for the improvement of thedataset. They used the KDD dataset and resultedshown dynamic growth in the result of the IDS.

**DISADVANTAGES OF EXISTING SYSTEM:**

* The systems which work over the internet suffer fromvarious malicious activities. The major problem seenin this field is the intrusion in the system for violatingthe information.
* Existing results state that there maybe some improvements to be done on terms ofaccuracy and the detection rates and the false alarmrate. Some other techniques can replace previouslyapplied techniques such as SVM and Naïve Bayes.Also, the study states that the dataset can beimproved by using some methods over it. To improvethe quality of the input to the proposed system.

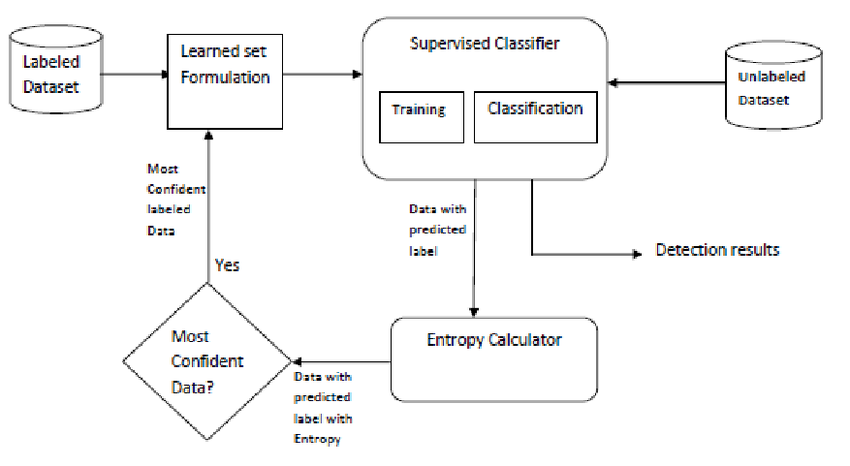
**PROPOSED SYSTEM:**

* The intrusion detection system works for theimprovement of the system, which is affected by theintruders. This system can do the detection of theintruders. The proposed system tries to eliminate theexisting problems related to the previous work. Theproposed system consists of the two methods that areprincipal component analysis, and the other one is therandom forest.
* The principal component analysis is used for thereduction of the dimension of the dataset; by thismethod, the dataset quality will be improved as thedataset may contain the correct attributes. After this,the random forest algorithm will be applied for thedetection of the intruders, which provide both thedetection rate and the false alarm rate in an improvedmanner as compared to SVM.

**ADVANTAGES OF PROPOSED SYSTEM:**

* The error rate found in our proposedapproach is very low as of .21%.
* As well, theaccuracy obtained is much higher than previousalgorithms.
* Also, the time taken for the performanceis less than other algorithms.

**SYSTEM ARCHITECTURE:**



**SYSTEM REQUIREMENTS:**

**HARDWARE REQUIREMENTS:**

* System : Pentium i3 Processor.
* Hard Disk : 500 GB.
* Monitor : 15’’ LED
* Input Devices : Keyboard, Mouse
* Ram : 2 GB

**SOFTWARE REQUIREMENTS:**

* Operating system : Windows 10.
* Coding Language : Python

**REFERENCE:**

SubhashWaskle, LokeshParashar, Upendra Singh, “Intrusion Detection System Using PCA withRandom Forest Approach”, IEEE CONFERENCE 2020.