Intrusion Detection System Using Deep Learning

Prof. Smita Jawale
Computer Engineering department
VCET College
Vasai, India
smita.jawale@vcet.edu.in

Shruti Sankhe
Computer Engineering department
VCET College
Vasai, India
shrutisankhe4@gmail.com

Rutuja Parab
Computer Engineering department
VCET College
Vasai, India
Rutujarockstarparab555@gmail.com

Nehab Shaikh
Computer Engineering department
VCET College
Vasai, India
nehabkalim@gmail.com

***Abstract* — these days, Intrusion Detection System (IDS) has become an important layer altogether the most recent Ict entity thanks to associate urge towards cyber safety within the daily world. Reason together with uncertainty within the finding of varieties of attacks and doubled the quality of refined cyber-attacks, IDS needs the necessity of integration of Deep Neural Network (DNNs). in this paper, DNNs are accustomed foretell the attacks on Network Intrusion Detection Entity (N-IDS). A DNN with zero.1 rate of learning is applied and is last thousand variety of epochs and KDDCup-'99' dataset has been used for coaching and benchmarking the network.**
***Index Terms*—Intrusion detection, deep neural networks, machine learning, deep learning.**

Index Terms—Intrusion detection, deep neural networks, machine learning, deep learning

Introduction

In the contemporary world, the fast technological advancements have inspired each organization to adopt the mixing of data and communication technology (ICT). Therefore making Associate in nursing atmosphere wherever each action is routed through that system creating the organization vulnerable if the safety of the ICT system is compromised. Therefore, this entail a multilayered finding on and protection theme that may handle actually novel attacks on the system similarly as in a position autonomously adapt to the new data. Intrusion find Systems (IDSs) are a spread of cybersecurity based mostly technology at the start developed to detect vulnerabilities and exploits against a target host. The only real use of the IDS is to find threats. Thus it's situated out-of-band on the infrastructure of the network and isn't within the actual time period communication passage between the sender and receiver of information.

I. Literature Review

1. Deep Neural Network (DNN)

Each layer applies a nonlinear transformation onto its input and creates a applied mathematics model as output from what it learns. In straightforward terms, the input layer is received by the input layer and passed onto the primary hidden layer. These hidden layers perform mathematical computations on our inputs. one in every of the challenges in making neural networks is deciding the hidden layers' count and also the count of the neurons for every layer.

2. Neural Network Based Intrusion Detection

In, category modification model consists of two-stages: i) P-rules stage to predict the presence of the class, and ii) N-rules stage to predict the absence of the category. This performed well compared with the same KDDCup ninety nine results apart from the user-to-root ('U2R') class. In, the importance of feature connectedness analysis was investigated for IDS with the foremost wide used dataset, KDDCup 99. for every feature they were ready to specific.

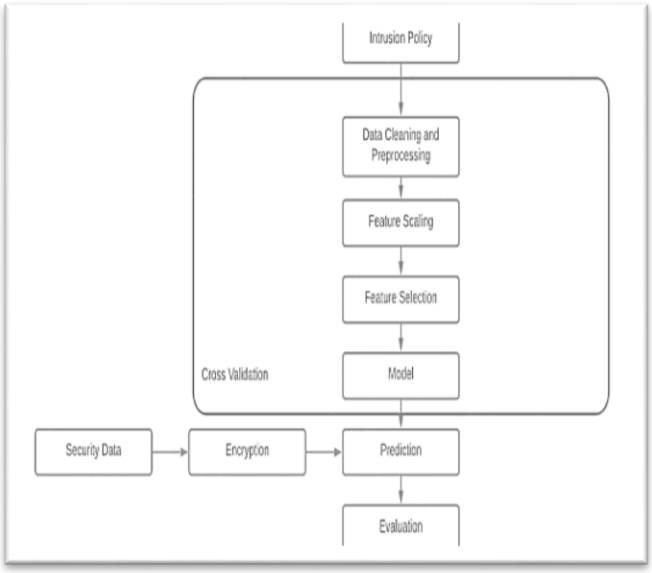
II. Analysis

2.1 Types Of Attacks

- Denial-of-Service-Attack(DoS):Intrusion wherever a per- son aims to form a bunch inaccessible to its actual purpose by in short or typically for good disrupting services by flooding the target machine with huge amounts of requests and therefore overloading the host.
- User-to-Root-Attack (U2R): A class of normally used maneuver by the culprit begins by attempting to achieve access to a user's pre-existing access and exploiting the holes to get root management.
- Remote-to-Local-Attack (R2L): The intrusion within which the wrongdoer will send knowledge packets to the target however has no user account thereon machine itself, tries to use one vulnerability to get native access cloaking themselves because the existing user of the target machine.
- Probing-Attack: the {sor|the kind} within which the culprit tries to assemble data concerning the computers of the network and therefore the final aim for doing so is to urge past the firewall and gaining root access.

III. Design

3.1 Activity Diagram Of IDS



IV. Proposed Architecture

An overview of projected DNNs design for all use cases is showcased in. This includes of a hidden-layer connote of five ANd an output-layer. The input-layer consists of forty one neurons. The neurons in input-layer to hidden-layer and hidden to output-layer square measure associated utterly. Back-propagation mechanism is employed to direct the DNN networks. The projected network is formed of totally associated layers, bias layers and dropout layers to form the network a lot of strong.

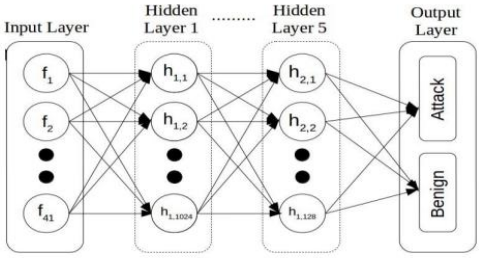
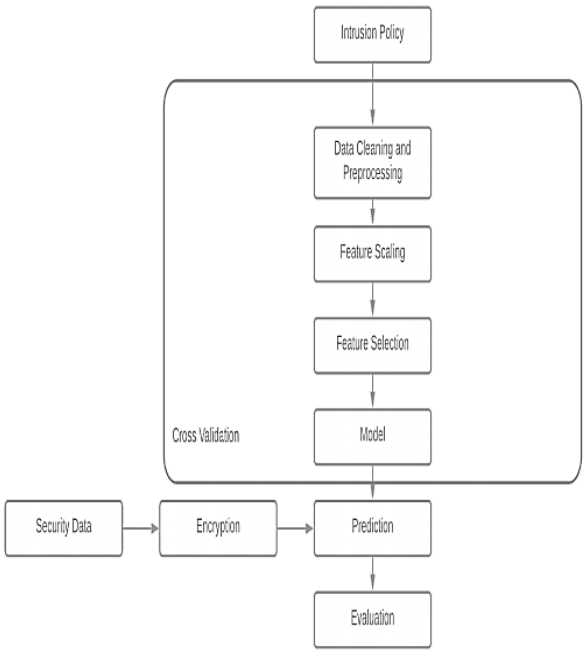


Fig 1 Proposed Architecture

5.1 Network Structure Information

- Input and hidden layers: This layer consists of forty one neurons. These ar then fed into the hidden layers. Hidden layers use ReLU as the non-linear activation operate. Then weights ar enclosed to feed them send to succeeding hidden layer. The somatic cell connote in every hidden layer is cut steady from the first to the output to form the outputs a lot of correct and at identical time reducing the recursive price.
- Output layer and classification: The out layer consists solely of 2 neurons Attack and Benign. Since the 1024 neurons from the preceding layer should be regenerate into simply a pair of neurons, a sigmoid activation operate is employed. because of the character of the sigmoid operate , I t reverts solely 2 outputs, second favouring the binary classification that was meant during this paper.

V. Implementation



Data Cleaning And Preprocessing:

The duplicates have already been removed as NSL KDD dataset is already standardised. Preprocessing operation is completed on the dataset because the dataset contains numerical and non-numerical values. this may convert all the explicit choices to their corresponding binary choices out of that one are active at a time. The dataset is then divided into four components as same by the attacks (U2R,Probe,DOS,R2L) that needed to be classified.

Features Scaling:

Featuring scaling is performed to steer distinct of choices that have huge values as this may bear on the ultimate result. Classic pulse counter is used to compose this operation. In Classic pulse counter the common for a possibility is calculated so the intend is deducted from the common charge of the choice and also the result's divided by the quality Deviation. the quality deviation are one once every feature is scaled.

Feature Selection And Model :

It's the method within which area unit immaterial and supererogatory choices area unit eliminated with tokenish databases loss. Subsets of the choices area unit choosen that totally serves all the choices within the dataset in terms of accuracy and alternative metrics. It's to boot doable that over there's a correlation between choices once a giant range of choices area unit gift. choice alternative additionally supports to induce obviate this downside.

VI. Result And Evaluation

We calculate four metrics on the premise of that we tend to assess the performance of our Intrusion Detection System. Given below is that the confusion matrix that may be a assortment of 4 values: TP, TN, FP, FN.

	ACTUAL VALUES		
PREDICTED VALUES		Actual: NO	Actual: YES
	Predicted: NO	True Negative	False Negative
	Predicted: YES	False Positive	True Positive

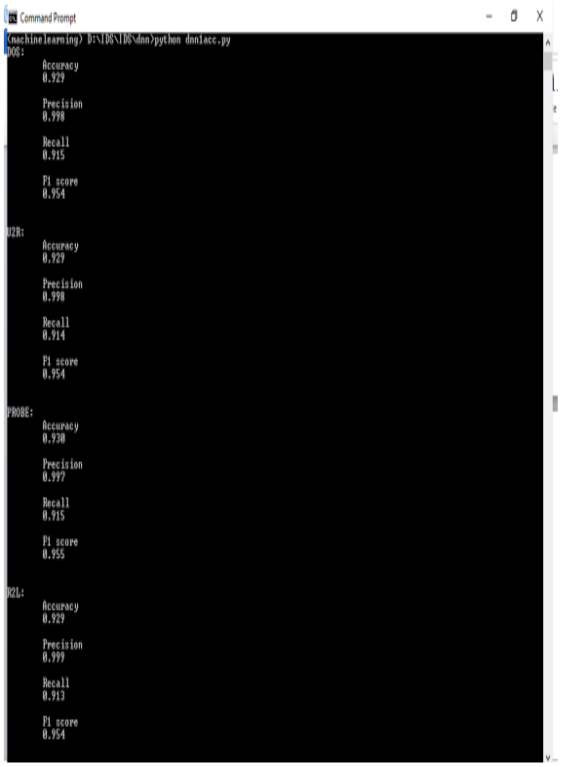
METRICS	FORMULA
Accuracy Rate	TP+TN/TP+FP+FN+TN
Recall(Detection Rate)	TP/TP+FN
Precision	TP/TP+FP
F-1 Score	2*(Recall * Precision) / (Recall + Precision)

VII. Results

The KDDCup-'99' dataset was fed into classical mil algorithm furthermore DNNs of varied hidden layers. once the coaching is completed, all models were compared for f1-score, accuracy, recall and exactness

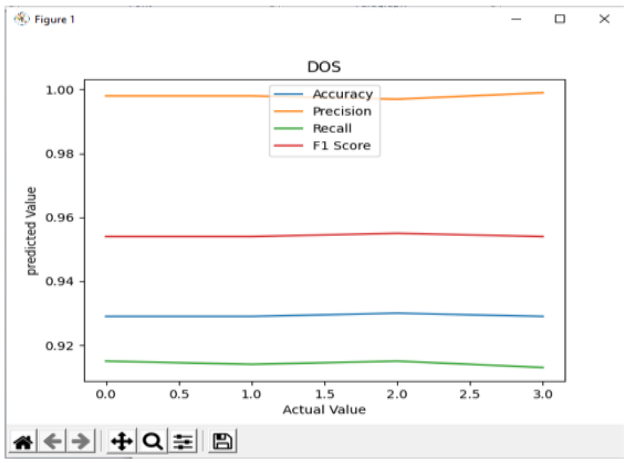
with the check dataset.

Cmd Output :

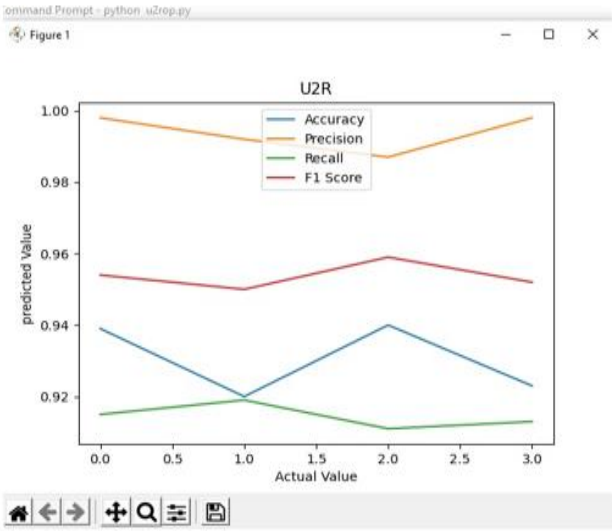


Histogram :

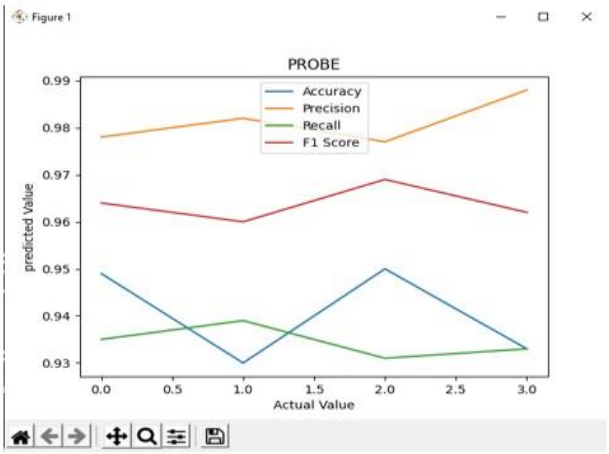
DOS Output:



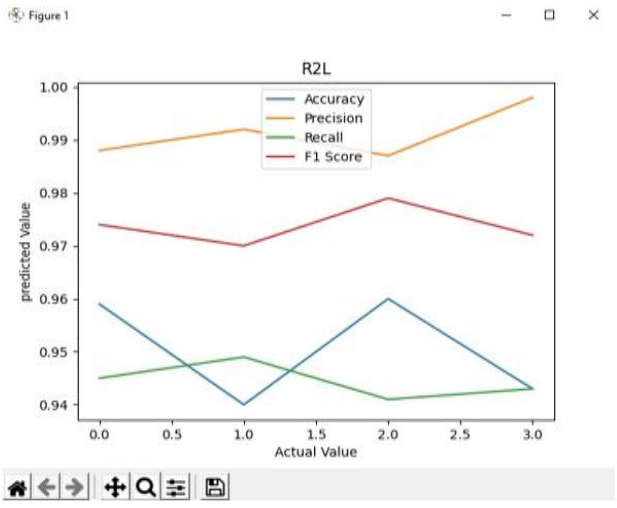
U2R Output:



PROBE Output:



R2L Output:



VIII. Conclusion

This paper has in an elaborate way recapitulated the utility of DNNs in IDS comprehensively. For the aim of reference, alternative classical milliliter algorithms are accounted and compared against the results of DNN. The in public obtainable KDDCup-'99' dataset has been primarily used because the bench-marking tool for the study, through that the prevalence of the DNN over the opposite compared algorithms are documented clearly. For any refinement of the algorithmic rule, this paper takes under consideration of DNNs with completely different counts of hidden layers and it absolutely was terminated that a DNN with three layers has been tested to be effective and correct of all.

IX. Reference

[1] S.Niksefat, P.Kaghazgaran and B.Sadeghiyan "Privacy issues in intrusion detection systems: A taxonomy, survey and future directions". Comput. Sci. Rev., 25, 69–78,2017.

[2] D. Archer, L. Chen, J. H. Cheon, R. Gilad-Bachrach, R. A. Hallman, Z. Huang, X. Jiang, R. Kumaresan, B. A. Malin, H. Sofia, Y. Song, and S. Wang, "Applications of homomorphic encryption," Homomorphic Encryption.org, Redmond WA, Tech. Rep., July 2017.

[3] B. Pfahringer. "Winning the KDD99 classification cup: Bagged boosting".
SIGKDD explorations newsletter, vol. 1, pp. 6566, 2000. DOI
http://dx.doi.org/10. 1145/846183.846200.

[4] Sarathiel Chaipa ; Mariki M Eloff ; Mariki M Eloff, "Towards the development of an Effective Intrusion Based Detection Model", 2017 Information Security for South Africa (ISSA), Johannesburg, South Africa, 16-17 Aug. 2017, Pretoria, IEEE.