INTRUSION DETECTION TECHNIQUES OF THE COMPUTER NETWORKS USING MACHINE LEARNING

Nilamadhab Mishra[1]*, Sarojananda Mishra[2]

Abstract
A network intrusion is an unauthorized operation of a computer network. The goal of a network access program is to protect computer networks from unauthorized users, including internal users. Create a local network discovery. This is a predicted form that distinguishes between "high quality" or typical associations and "terrible" associations, sometimes called intruders or attacks. The purpose was to evaluate accessibility results. We also concentrated on machine learning-based classification to facilitate acquires greatest training and testing, to access our strategy for using currently available technologies. To generate various classification models, used varieties machine-learning based techniques and comparing each other for detecting best fit model for the computer networks with respect to time and accuracy. Based on comparisons with six different machine learning algorithms used to categories attacks, it is feasible to determine the originality of the proposed IDT by the fact that it utilizes the best machine learning method possible to fit into a high performance IDT.

Keywords: Intrusion Detection Techniques, Machine-Learning, Computer Networks, Classification Approach, Performance Measure.

[1*]Ph.D Scholar, Engineering, Biju Patnaik University of Technology, Rourkela, Odisha, India,
[2]Professor, Indira Gandhi Institute of Technology, Sarang , Dhenkanal, Odisha, India

***Corresponding Author: -** Nilamadhab Mishra
*Ph.D Scholar, Engineering, Biju Patnaik University of Technology, Rourkela, Odisha, India,
E-Mail: nilamadhab76@gmail.com

DOI: - 10.48047/ecb/2023.12.si5a.054

1. Introduction

Intrusion detection technology (IDT) is a control technology, either physical or programmatic, that examines data in a network or fabric to identify intruders [1]. IDS identify threats using three techniques: Asshole-based detection, uncontrolled detection, and signature-based detection. Identifies known attacks using signature-based detection by examining signatures.

A good technique for finding known attacks recorded in the IDS database. As a result, it was often believed that an effective identification attempt had been made, or that a known attack had taken place. Newer forms of abuse, on the other hand, are unrecognizable due to the absence of handwriting.

Data information is frequently refreshed to improve performance levels. This problem is solved by using uncontrolled selective detection based on the most recent customer and previous profiles.

Identification of potentially destructive behavior. Anomaly-based detection is effective against unconfirmed he 0'Day threats and all system updates. However, this strategy has many conceptual advantages [2].

A computer program called Access Login uses machine learning to detect network access. IDS protects against unauthorized access by users, including insiders, and detects networks or systems that exhibit malicious activity. The purpose of this research is to develop a form of intruder prediction that can distinguish between normal and attack connections.

A classification problem is an attack detection problem that reveals whether a data packet is an attack type or a normal type. As such, IDT was implemented using various machine learning (ML) techniques.

Here, the authors implement various ML algorithms on the approved Knowledge-Discovery (KDD) dataset, including attack types such as Daniel-of-Service (dos), r2l, u2r, and probe [5].

1.1 Literature Discussion

IDTs based literature is abounding with recent machine learning methods. Many proposed IDTs models found in classical machine learning methods which give low accuracy as well as the depends on the manual process to design the traffic features.

In this paper, the author brings the effective IDS using DL. Collecting data from different standard datasets which contains different type of the attacks.

Then the data can be processed to eliminate the anomalies using the removal of missing value and technique of the normalization. Feature extraction using auto encoder (AE), removing timestamps from attack using Random Forest (RF) [3].

The Author's proposed a hybrid model which combines the machine learning and deep learning to improve the detection rate . Here for data balancing SMOTE and for feature selection XGBoost have implemented to develop a novel, dependable and effective network intrusion detection system with Machine Learning and Deep Learning [4].

In order to detect all types of attacks, including user2root (u2r) and remote2local (r2l) attacks, the authors of [5] acknowledged that a single ML classifier is not helpful. Instead, they suggested using signature-based IDTs to detect these attacks.

As a result, the proposed IDT employs a two-layered hybrid strategy in which Naive-Bayes identify Daniel-of-Source and PROBE in layer one and SVM detection of u2r and r2l in layer two accomplish the desired objective.

Objective of the intrusion detection system is to manage the network performance and detect the abnormalities over the network. The author's proposed a model for intrusion detection and classification using machine learning techniques. Here used Konstanz information Minor (KNIME) to refined the dataset and for better performance and comparative study three classifiers are used like SVM, RProp and Decision Tree [6].

The author's modeled an intrusion detection system using six different machine learning algorithms to classify the attack and normal type. The performances have been analyzed using different performance measure and found the best fit with respect to accuracy and time [7].

For the effective data processing, detection of harmful behavior and control the identification of attack author's proposed machine learning techniques. Here four types of the attacks

predicted with the implementation of the ensemble model to enhance the performance using AdaBoost and logistic regression.
Then it is compared to other work [8].
The author's in this work evaluate the performance of fifteen different machine learning techniques out of which five are selected on the basis of maximum accuracy and minimum errors in WEKA. The simulation can be done using 10 fold cross validation the best ML algorithm selected.
The main objective is to detect the effective and perfect machine learning algorithm which controlled the network intrusion in a suitable manner. Out of fifteen different ML algorithm Random Tree poses more accuracy on high dimensional data [9].
In this work author's proposed the Neighbor Distance Variance classifier for the prediction purposes. It is a binary class predictor which implements the concept of the variance of the distance between the objects.
Used KDD CUP-99 dataset to examine the NNDV and compared the predicted accuracy of NNDV with the KNN or K-Nearest Neighbor classifier. KNN is an efficient classifier, but here only considered its binary aspect.
The outcome is manageable to show that NNDV is comparable to KNN. And also compared the accuracy results of different cross-validation techniques used such as 2-fold, 5-fold, 10-fold, and exclude-one-out on the NNDV for the KDD CUP-99 dataset. The parameters of the algorithm can be detected with the help of the Cross-validation results [10].
The study proposed by the author's to evaluate feature extractors including the image filter and shift the learning models like VGG-16 and DenseNet. Then different ML algorithms implemented for feature extraction. This work presented the evaluation of the combined models by using IEEE dataport databases [11].
The author's proposed an intrusion detection system utilizing the machine learning algorithms. Different machine learning algorithms such as Support Vector Machine, J48, Random Forest, and Naive Bayes with binary and multiclass classification have been implemented. Here random forest performs well to detect the intrusion [12].
The author introduced an approach based on HOA for the IDS. Here quantum computing and HOA combining improve the behavioral characteristics. The proposed algorithm MQBHOA have adopted for intrusion detection of the computer networks which itself is a multi objective optimization problem. For classification KNN is applied [13].
Anomaly based intrusion detection have proposed by the author's which recognized all type of the attacks with better accuracy. This work based on the imbalanced data which is processed by random over-sampling algorithm and again optimized by different high end optimizers of deep neural network[14].
In [15][16] the authors propose an IDT model that combines the mechanism of attention with Bidirectional long short-term memory (BLSTM), which uses the BLSTM method to automatically extract traffic data from network flow. The adopted artificial intelligence classifier uses unprocessed data as its input, not features that were manually designed.
The authors of this learning strategy did not address the tuning of CNN's parameters. Additionally, the ML method used was not tested for its capability. Because it is compressed, the proposed method does not validate unknown malware traffic, which indicates the scope of subsequent work.

Table 1: Snapshot of Literature Survey

References	Year	Algorithm	Main Contribution	Field
[2]	2020	Naive-bayes, J-48, and Random -forest	For the design and implementation of Random-forest, work well in IDT.	Machine Learning(ML)
[3]	2023	Auto encoder, Random Forest	Effective implementation of Deep Learning to detect attacks	Deep Learning
[4]	2023	SMOTE, XGBoost	Effective network intrusion detection system	Machine Learning and Deep Learning
[5]	2021	Naive-Bayes and SVM classifier	A double-layered hybrid approach (DLHA) was proposed by the authors.	Machine-Learning(ML)
[6]	2022	SVM, RProp and Decision Tree	To manage the network performance and detect the abnormalities over the network.	Machine Learning
[7]	2023	Naive-Bayes, DT, RF, SVM, LR, GD	The performances have been analyzed with respect to accuracy and time	Machine Learning
[8]	2023	AdaBoost and logistic regression	Implementation of the ensemble model to enhance the performance.	Machine-Learning(ML)

[9]	2016	10 fold cross validation, Random Tree	Detect the effective and perfect machine learning algorithm which controlled the network intrusion in a suitable manner	Machine-Learning(ML)
[10]	2023	KNN	Parameters of the algorithm can be detected with the help of the Cross-validation results.	Machine-Learning(ML)
[11]	2023	Shift the learning models like VGG-16 and DenseNet	Evaluate feature extractors including the image filter and	Machine-Learning(ML)
[12]	2022	Support Vector Machine, J48, Random Forest, and Naive Bayes	Better Intrusion detection by Random Forest.	Machine-Learning(ML)
[13]	2023	algorithm MQBHOA with KNN	Quantum computing and HOA combining improve the behavioral characteristics	Quantum computing and Machine Learning
[14]	2023	Random over sampling and DL	All type of the attacks with better accuracy.	Deep Learning
[15]	2020	LSTM , CNN	To detect each attack type LSTM and CNN models are proposed.	Deep Learning(DL)
[16]	2022	ADASYN and CNN	For better performance classification proposed a model DLNID.	Deep Learning(DL)
[18]	2020	ML and DL	AI-based NIDS	Machine Learning(ML)/ Deep Learning(DL)

2. Background

Future forecast exactness enhances without premeditated using machine learning which utilize chronological input data [19].

The machine-learning-techniques are commonly used in recommendation engines. Also different techniques are used for scam-recognition; spamming-filtration; detection of malware risk; Business-process-automation (BPA); as well as analytical maintenances are general relevances. [20]

2.1Machine-Learning-Types:

Conventional machine-learning is frequently categorized as the practice, through this algorithm, increase the accurateness of its predictions. Four main approaches are supervised, unsupervised, semi supervised, and reinforcement learning methods. By utilizing the data, data analysts want to foresee the algorithm that they choose.

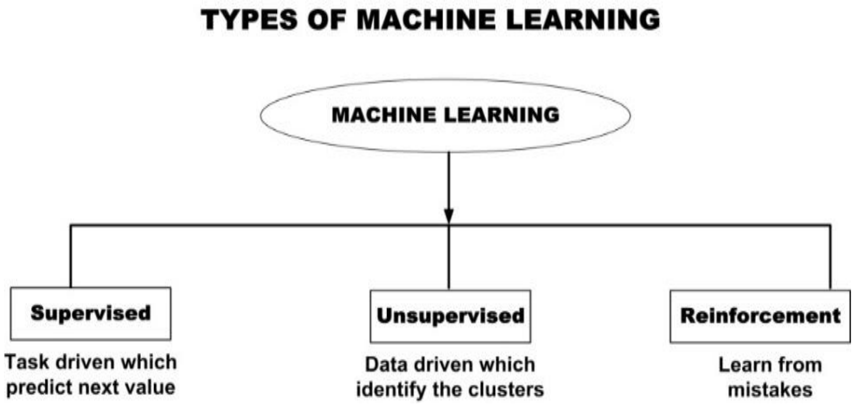

Figure 1. Machine learning types

2.2 Various Classification Algorithms:

Guassian-Naïve-Bayes Algorithm(NBA):

The method is implemented to establish a classification model with only numerical values and classify both documents and text. It is very simple to train and use also can easily predict classes. It is a given that class has no bearing on features. Applications for the naive bayes algorithm include reaction study, recommender system, and spam filter [21].

Decision Tree Algorithm(DTA):

DTA is the fundamental of supervised-learning method which uses a series of decisions, for classification and prediction of data (rules). The model is organized into nodes, branches, and leaves like a tree.

Every node stands for a property or an attribute. The branch stands for a choice or a protocol, where each leaf denotes a potential outcome or a name of the class. The DTA method automatic chooses the good qualities for constructing tree, and henceforth prunes the tree to get rid of needless branches in order to reduce over-fitting.

Random Forest Algorithm (RFA):
RFA is an established machine learning algorithm that falls under the supervised approach category. It can be used to solve problems related to classification and regression. Adapted the model's performance capability by combining various classifiers to tackle a problem using ensemble learning.
"Random-forest(RF) algorithm incorporates multiple decision-trees on different subsets of the given dataset and takes the average to boost the projected accuracy of that dataset," as the name suggests, is the function of the "random-forest" algorithm.
The forecast from each decision tree and the majority prediction of votes are then used by the random forest algorithm to predict the final outcome.

Algorithm for Support Vector Machines (SVMA):
The idea of the hyper plane along with greatest partitioned of margin in nth-dimensional attributes spaces serve as the foundation for the supervised machine learning method known as SVM. It is capable of handling both linear and nonlinear issues. Nonlinear problems are resolved with the function of kernel. The objective is to convert a vector of low dimensional inputs into a high dimensional features by implementing the kernel function. The ideal maximize marginal-hyper-plane is then discovered using the support vectors and acts as a decision boundary. The accuracy and efficiency of NIDS can be increased by using the SVM algorithm to accurately forecast the normal and dangerous classifications[5]. Extreme vectors and points can be selected by SVM that help to create hyper-plane. Support-vector, which are symbolize these excessive instances on the basis of the support vector -method.

Logistic Regression (LR):
Logical regression, a supervised classification algorithm, only accepts distinct value as input and generates a regression-based-model that foretells whether known pieces of information have a likelihood of being 1 or zero (0).

These values can refer to any of the classifications used to group data. Logistic regression can be used rapidly to identify the factors that will work well when classifying observations using various sources of data.

Gradient Descent Algorithm (GDA):
The most frequent optimization method is gradient-descent, which is utilized in deep-learning and machine-learning algorithm. It is a forwarded optimization technique which is used to consider the first derivative when changing the parameters.

In every repetition, we have to change the parameter in the reverse path of the goal function J (w) gradient, where the gradient denotes the sharpest ascending direction. To achieve the local-minimum, take the size of each step depending on the rate of learning.

As a result of which, need to continue or to move downward until reach to a local minimum. The important purpose of a Gradient-descent method is to repeat minimizing cost-function. It carries out the following steps repeatedly to reach the goal.

3. Simulation and Result:

3.1 **Workflow Diagram:**

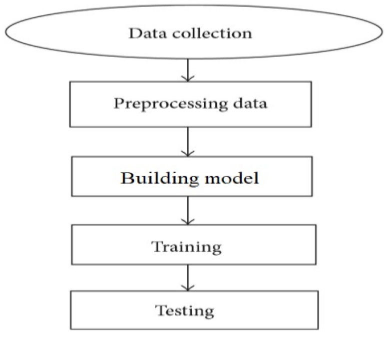

Figure 2. Workflow diagram of the data analysis

3.2Preprocessing of Data :

List of features reading from "Kddcup.names" file by importing concern file. Adding new column to the dataset as 'target' through which find out 42-features. Reading of 'Attack_Types' files hown in the Table- 2:

Table 2. List of attack _types

CLASS	ATTACK TYPE
dos	Disconnect of the network service(pod, Neptune, Smurf)
r2l	Guessing of password(multihop, Phf, Warezclient)
u2r	Over flow of the buffer (loadmodule, Rootkit, Perl)
Probing	Scanning of port(portsweep, Nmap, Satan)

Creating of the dictionary using attack types.The reading and features of the [attack-type] dataset (["kddcup.data_10_percent.gz") have been added to the training_dataset. This dataset contains five distinct attack type features: dos, normal, Probe, u2r, and r2l. Determining the data type of each feature and shaping the data frame. Finding missing values but here no missing value have found, then we go for further step. Categorical Features have been found out: ['service ', 'flag ', 'protocol-type'] . Finding correlated variables by the implementation of heat-map and exempted them for scrutiny. Mapping of feature – Applied feature mapping on 'protocol-type' & 'flag '. Removed unrelated facial appearance for instance 'service' before modeling.

3.3 Modeling:

Libraries importing and dataset splitting. Dataset divided as [494021, 31]. Training and Testing data splitting which available in Table-3.

Table 3. Training and Testing data Splitting

X_train_data	***X_test_data***
[330994, 30]	[163027, 30]
y_train_data	**y_test_data**
[330994, 1]	[163027,1]

Using various machine-learning classification algorithms like: We obtain the following trained and tested results from the Naive Bayes Algorithm (NBA), Decision Tree Algorithm (DTA), Random Forest Algorithm (RFA), Support Vector Classifier Algorithm (SVCA), Logistic Regression Algorithm (LRA), and Gradient Descent Algorithm (GDA) represent in Table-4:

Table 4. List of Score Training and Testing

Algorithim	***Training***	***Testing***
NBA	87.951	87.903
DTA	99.058	99.052
RFA	99.997	99.964
SVCA	99.875	99.879
LRA	99.352	99.352
GDA	99.793	99.771

4. Result Analysis

i) The train and test accuracy of each model analysis using Table:6 is given in Figure 10 :

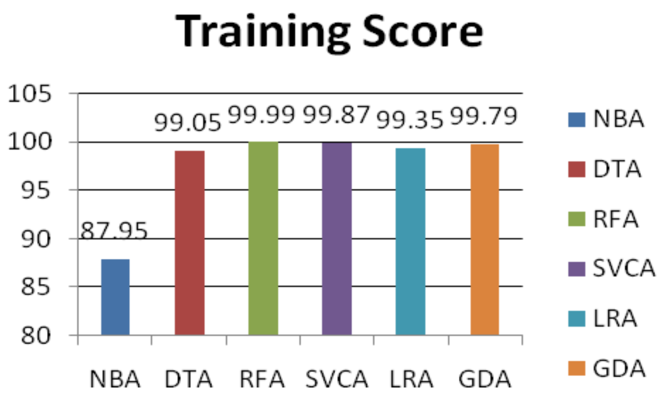

Figure 3. Training accuracy analysis

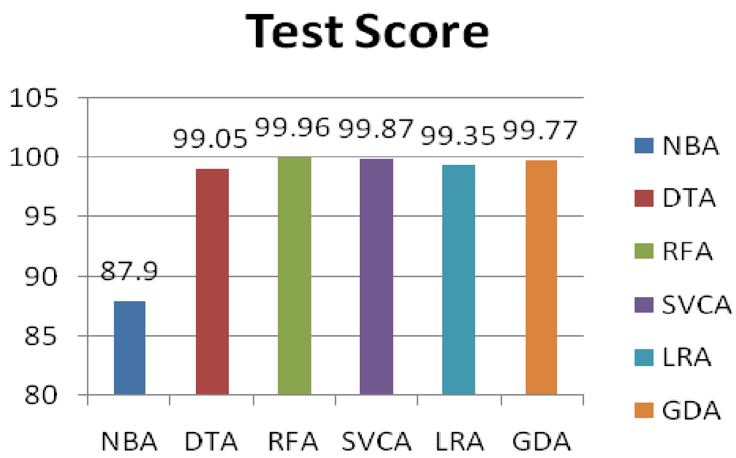

Figure 4. Testing accuracy Analysis.

ii) **T**raining and testing time analysis :
Different algorithms should be implemented on train and test dataset and the time required to train and test data results are given in following figures:

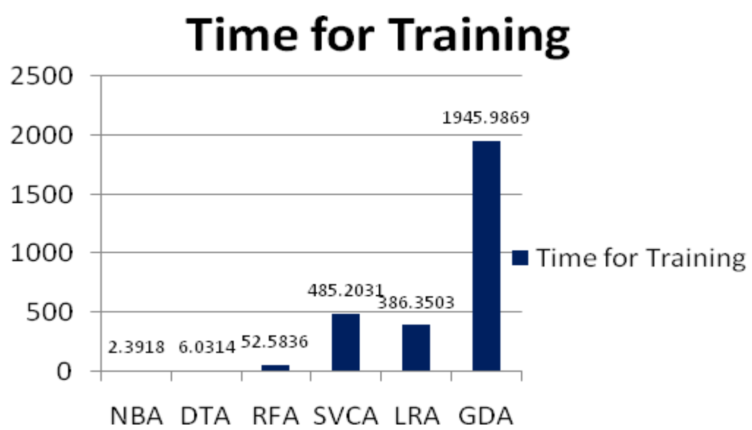

Figure 5. Analysis of the training time

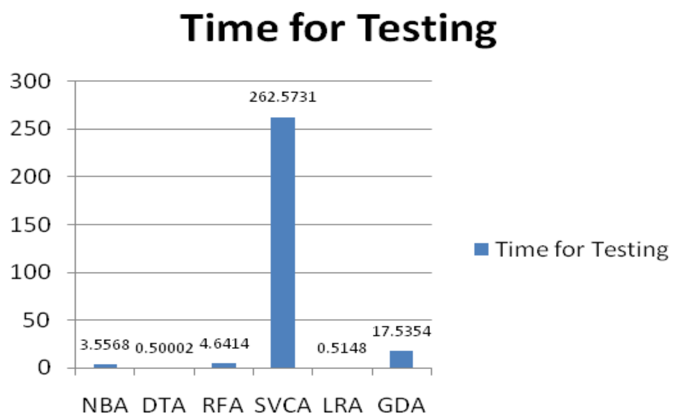

Figure 6. Analysis of the testing time

5. Classification Report using different Machine Learning Techniques:

Table 5. Classification report using Gaussian Naive Bayes

	precision	*recall*	*f1-score*	*support*
DoS	1.00	0.94	0.97	389717
R2L	0.03	0.42	0.05	125
U2R	0.01	0.83	0.03	6
Probe	0.02	0.99	0.04	456
Accuracy			0.94	390304
Macro-Avg	0.21	0.64	0.22	390304
Weighted-Avg	1.00	0.94	0.97	390304

Table 6. Classification report using Decision Tree

	precision	*recall*	*f1-score*	*support*
DoS	1.00	0.94	0.97	389717
R2L	0.64	0.84	0.72	125
U2R	1.00	0.50	0.67	6
Probe	0.02	1.00	0.04	456
Accuracy			0.95	390304
Macro-Avg	0.53	0.66	0.48	390304
Weighted-Avg	1.00	0.95	0.97	390304

Table 7. Classification report using Random Forest

	precision	**recall**	**f1-score**	**support**
DoS	1.00	1.00	1.00	389717
R2L	0.92	0.99	0.95	125
U2R	0.50	0.83	0.62	6
Probe	0.69	0.99	0.81	456
Accuracy			1.00	390304
Macro-Avg	0.62	0.76	0.68	390304
Weighted-Avg	1.00	1.00	1.00	390304

Table 8. using Support Vector Classifier classification report

	precision	**recall**	**f1-score**	**support**
DoS	1.00	0.99	1.00	389717
R2L	0.76	0.93	0.84	125
U2R	1.00	0.50	0.67	6
Probe	0.51	0.98	0.67	456
Accuracy			0.99	390304
Macro-Avg	0.66	0.68	0.63	390304
Weighted-Avg	1.00	0.99	1.00	390304

Table 9. Using Logistic Regression classification report

	precision	recall	f1-score	support
DoS	1.00	0.99	1.00	389717
R2L	0.74	0.90	0.81	125
U2R	1.00	0.50	0.67	6
Probe	0.53	0.96	0.68	456
Accuracy			0.99	390304
Macro-Avg	0.65	0.67	0.63	390304
Weighted-Avg	1.00	0.99	1.00	390304

Table 10. Using Gradient-descent classification report

	precision	recall	f1-score	support
DoS	1.00	1.00	1.00	389717
R2L	0.97	0.99	0.98	125
U2R	0.50	0.67	0.57	6
Probe	0.74	0.99	0.84	456
Accuracy			1.00	390304
Macro-Avg	0.64	0.73	0.68	390304
Weighted-Avg	1.00	1.00	1.00	390304

6. Performance Comparison:

Table-11 Performance comparison with respect to time

Overall performance comparison of proposed methods							
Type	***Training time***	***Testing Time***	***ACC (%)***	***CLASS***	***Pre (%)***	***Rec (%)***	***F1 (%)***
NB	2.3918	3.5568	0.94	DoS	1	0.94	0.97
				R2L	0.03	0.42	0.05
				U2R	0.01	0.83	0.03
				Probe	0.02	0.99	0.04
DT	6.0314	0.50002	0.95	DoS	1	0.94	0.97
				R2L	0.64	0.84	0.72
				U2R	1	0.5	0.67
				Probe	0.02	1	0.04
RF	52.5836	4.6414	1	DoS	1	1	1
				R2L	0.92	0.99	0.95
				U2R	0.5	0.83	0.62
				Probe	0.69	0.99	0.81
SVC	485.2031	262.5731	0.99	DoS	1	0.99	1
				R2L	0.76	0.93	0.84
				U2R	1	0.5	0.67
				Probe	0.51	0.98	0.67
LR	386.3503	0.5148	0.99	DoS	1	0.99	1
				R2L	0.74	0.9	0.81
				U2R	1	0.5	0.67
				Probe	0.53	0.96	0.68
GD	1945.9869	17.5354	1	DoS	1	1	1
				R2L	0.97	0.99	0.98
				U2R	0.5	0.67	0.57
				Probe	0.74	0.99	0.84

Table-12 Performance comparison with other research work

Overall Performance comparison to other research work					
Type of IDS	***ACC (%)***	***CLASS***	***Pre (%)***	***Rec (%)***	***F1 (%)***
Proposed IDT with DT implementation	0.95	DoS R2L U2R Probe	1 0.64 1 0.02	0.94 0.84 0.5 1	0.97 0.72 0.67 0.04
LNID [16]	0.90	u2r	0.86	0.93	0.89
DLHA [5]	0.87	r2l, u2r	0.88	0.90	0.89
BAT-MC[15]	0.84	dos,normal,probe,r2l,u2r			
Autoencoder [28]	0.84	Normal, DoS, R2L and Probe	0.87	0.80	0.819
CNN [29]	0.80	r2l, u2r			
Adaptive Ensemble [11]	0.85	Probe, R2L and U2R	0.86	0.86	0.85
GAR-Forest [30]	0.85	dos,normal,probe,r2l,u2r	0.87	0.85	0.85
CNN+BiLSTM [31]	0.83	dos,normal,probe,r2l,u2r	0.85	0.84	0.85
NB Tree [32]	0.82	dos,normal,probe,r2l,u2r			
SVM-IDS [33]	0.82	dos,normal,probe,r2l,u2r			

7. Conclusion

In this analysis of Intrusion Detection Techniques, the best possible machine learning algorithm for efficient high-performance IDT matching, six different machine learning algorithms were modeled to classify attack types, normal and bad. All classifiers were trained and tested using the KddCup dataset. The performance of classifiers is analyzed as different performance measures, such as evaluation of training and test results scores, training and testing schedules using different techniques, and also generating produce a classification report for each technique. According to training and testing, time and report scores are analyzed, it is found that Decision Tree Modeling (DTM) is one of the best data classification techniques implemented in this work. and assessed accuracy and time complexity according to the classification results with 95% Accuracy, better than other authors in their study.

Acknowledgments

This study has not received any specific grants from public, commercial, or nonprofit funding agencies. The authors would like to thank the editor, guest editors and anonymous reviewers for their comments to help improve the quality of this work.

Conflict of Interest

The authors confirm that there is no conflict of interest to claim for this publication.

References:

1. S. M. Othman, F. M. Ba-Alwi, N. T.

Alsohybe, and A. Y. Al-Hashida, “Intrusion detection model using machine learning algorithm on Big Data environment,” *J. Big Data*, vol. 5, no. 1, 2018, doi: 10.1186/s40537-018-0145-4.
2. C. Aishwarya*, N. Venkateswaran, T. Supriya, and V. Sreeja, “Intrusion Detection System using KDD Cup 99 Dataset,” *Int. J. Innov. Technol. Explor. Eng.*, vol. 4, no. 9, pp. 3169–3171, 2020, doi: 10.35940/ijitee.d2017.029420.
3. Liloja1* and Dr.P. Ranjana2 ," An Intrusion Detection System Using a Machine Learning Approach in IOT-based Smart Cities ", Journal of Internet Services and Information Security (JISIS), volume: 13, number: 1 (February), pp. 11-21 DOI: 10.58346/ JISIS.2023.I1.002,2023
4. Md. Alamin Talukdera , Khondokar Fida Hasanb , Md. Manowarul Islama , Md Ashraf Uddina , Arnisha Akhtera , Mohammad Abu Yousufc , Fares Alharbid , Mohammad Ali Moni," A Dependable Hybrid Machine Learning Model for Network Intrusion Detection ", Journal of Information Security and Applications on 7 December 2022. arXiv:2212.04546v2 [cs.CR] 27 Jan 2023.
5. T. Wisanwanichthan and M. Thammawichai, “A Double-Layered Hybrid Approach for Network Intrusion Detection System Using Combined Naive Bayes and SVM,” *IEEE Access*, vol. 9, pp. 138432–138450, 2021, doi: 10.1109/ACCESS.2021.3118573.
6. Ameera S. Jaradat, Malek M. Barhoush, Rawan Bani Easa," Network intrusion detection system: machine learning approach ", Journal of Electrical Engineering and Computer Science Vol. 25, No. 2, February 2022, pp. 1151~1158 ISSN: 2502-4752, DOI: 10.11591/ijeecs.v25.i2.pp1151-1158.
7. Nilamadhab Mishra1*, Sarojananda Mishra2," A Novel Intrusion Detection Techniques of the Computer Networks Using Machine Learning ", International Journal of Intelligent Systems and Applications in Engineering IJISAE, 2023, 11(5s), 247–260 | 247
8. Mukesh Kumar Yadav1 , Mahaiyo Ningshen," Enhancement of Intrusion Detection System using Machine Learning ", International Journal of Engineering Research & Technology (IJERT) http://www.ijert.org ISSN: 2278-0181 IJERTV12IS010058 Published by : www.ijert.org Vol. 12 Issue 01, January-2023
9. Prachi , " Usage of Machine Learning for Intrusion Detection in a Network", International Journal of Computer Networks and Applications (IJCNA) DOI: 10.22247/ ijcna/2016/41278 Volume 3, Issue 6, November – December (2016)
10. Krishna Gopal Sharma*, Yashpal Singh , " Predicting Intrusion in a Network Traffic Using Variance of Neighboring Object’s Distance ", I. J. Computer Network and Information Security, 2023, 2, 73-84 Published Online on April 8, 2023 by MECS Press (http://www.mecs-press.org/) DOI: 10.5815/ijcnis.2023.02.06
11. Dhiaa Musleh 1 , Meera Alotaibi 1 , Fahd Alhaidari 2 , Atta Rahman 1,* and Rami M. Mohammad," Intrusion Detection System Using Feature Extraction with Machine Learning Algorithms in IoT ", J. Sens. Actuator Netw. 2023, 12, 29. https://doi.org/10.3390/jsan12020029 https://www.mdpi.com/journal/jsan
12. Yasmeen S. Almutairi1 , Bader Alhazmi1 , Amr A. Munshi1," Network Intrusion Detection Using Machine Learning Techniques ", Advances in Science and Technology Research Journal 2022, 16(3), 193–206 https://doi.org/10.12913/22998624/149934 ISSN 2299–8624
13. Reza Ghanbarzadeh1 · Ali Hosseinalipour2 · Ali Ghafari, " A novel network intrusion detection method based on metaheuristic optimisation algorithms ", Journal of Ambient Intelligence and Humanized Computing https://doi.org/10.1007/s12652-023-04571-3
14. Nilamadhab Mishra 1 , Sarojananda Mishra 2, Bhaskar Patnaik 3 ," A Novel Intrusion Detection System Based on Random Oversampling and Deep Neural Network ", Indian Journal of Computer Science and Engineering (IJCSE), e-ISSN : 0976-5166 p-ISSN : 2231-3850, doi: 10.21817/indjcse/2022/v13i6/221306136.
15. T. Su, H. Sun, J. Zhu, S. Wang, and Y. Li, “BAT: Deep Learning Methods on Network Intrusion Detection Using NSL-KDD Dataset,” *IEEE Access*, vol. 8, pp. 29575–29585, 2020, doi: 10.1109/ACCESS.2020.2972627.
16. Y. Fu, Y. Du, Z. Cao, Q. Li, and W. Xiang, “A Deep Learning Model for Network Intrusion Detection with Imbalanced Data,” *Electron.*, vol. 11, no. 898, pp. 1–13, 2022, doi: 10.3390/electronics11060898.
17. K. Vengatesan, A. Kumar, R. Naik, and D.

K. Verma, “Anomaly based novel intrusion detection system for network traffic reduction,” in *Proceedings of the International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), I-SMAC 2018*, 2018, no. August, pp. 688–690. doi: 10.1109/I-SMAC.2018.8653735.

18. Z. Ahmad, A. Shahid Khan, C. Wai Shiang, J. Abdullah, and F. Ahmad, “Network intrusion detection system: A systematic study of machine learning and deep learning approaches,” *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 1, pp. 1–29, 2021, doi: 10.1002/ett.4150.
19. S. Sah, “Machine Learning: A Review of Learning Types,” *ResearchGate*, no. July, 2020, doi: 10.20944/preprints202007.0230.v1.
20. E. Burns, “machine learning,” *Machine Learning*, 2021.
21. “machine-learning-algorithms.”https://www.javatpoint.com/machine-learning-algorithms
22. A. Abedinia, “Survey of the Decision Trees Algorithms (CART, C4.5, ID3) | by Aydin Abedinia | Medium,” 2019. https://medium.com/@abedinia.aydin/survey-of-the-decision-trees-algorithms-cart-c4-5-id3-97df842831cd
23. “machine-learning-decision-tree-classification-algorithm.” https://www.javatpoint.com/machine-learning-decision-tree-classification-algorithm
24. “Random Forest Algorithm,” 2018. https://www.javatpoint.com/machine-learning-random-forest-algorithm
25. “logistic-regression-in-machine-learning.” https://www.javatpoint.com/logistic-regressi- on-in-machine-learning
26. “Gradient Descent in Machine Learning.” https://www.javatpoint.com/gradient-descent-in-machine-learning
27. S. J. Stolfo, W. Fan, W. Lee, A. Prodromidis, and P. K. Chan, “KDD-CUP-99 Task Description,” 1999. http:// kdd. ics.uci.edu/databases/kddcup99/task.html
28. Ieracitano, C., Adeel, A., Morabito, F. C., & Hussain, A. (2020). A novel statistical analysis and autoencoder driven intelligent intrusion detection approach. Neurocomputing, 387, 51-62.
29. Gao, X., Shan, C., Hu, C., Niu, Z., & Liu, Z. (2019). An adaptive ensemble machine learning model for intrusion detection. IEEE Access, 7, 82512-82521.
30. Kanakarajan, N. K., & Muniasamy, K. (2016). Improving the accuracy of intrusion detection using gar-forest with feature selection. In Proceedings of the 4th International Conference on Frontiers in Intelligent Computing: Theory and Applications (FICTA) 2015 (pp. 539-547). Springer, New Delhi.
31. Jiang, K., Wang, W., Wang, A., & Wu, H. (2020). Network intrusion detection combined hybrid sampling with deep hierarchical network. IEEE Access, 8, 32464-32476.
32. Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009, July). A detailed analysis of the KDD CUP 99 data set. In 2009 IEEE symposium on computational intelligence for security and defense applications (pp. 1-6). Ieee.
33. Pervez, M. S., & Farid, D. M. (2014, December). Feature selection and intrusion classification in NSL-KDD cup 99 dataset employing SVMs. In The 8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA 2014) (pp. 1-6). IEEE.