IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 7, July 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

9940 572 462 | 6381 907 438 | ijircce@gmail.com | www.ijircce.com

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007014 |

Intrusion Detection of Imbalanced Network Traffic Based on Machine Learning and Deep Learning Technique

Aishwarya N Harigol, Dr. Md. Tajuddin

II M.Tech Student, Department of CSE, Dayananda Sagar College of Engineering, Bengaluru, Karnataka, India

Professor, Department of CSE, Dayananda Sagar College of Engineering, Bengaluru, Karnataka, India

ABSTRACT: Malicious cyberattacks may sometimes be camouflaged as enormous volumes of regular data when there is an imbalance in the network flow. NIDS detection accuracy and speed are difficult to achieve due to the high degree of stealth and opacity shown in cyberspace. Machine learning and deep learning are investigated in this research to discover intrusions in uneven network data. A special technique for difficult cluster sampling (DSSTE) is proposed to overcome the problem of category mismatch. First, separate the imbalanced training into the critical group and the normal group using the Nearest Neighbor (ENN) modification approach. The difficulty of zooming in and out of minority samples and subsequently creating additional samples to boost the minority population are recurring traits of a key group. The easier group is joined into a new training set, the minority into the tougher group, and the majority compact group into the harder group. The cluster is reduced using the KMeans technique.

KEYWORDS*:* IDS, asymmetric network traffic, deep learning, and the 2018 CSE-CIC-IDS.

I. INTRODUCTION

There has been a massive surge in the complexity and diversity of cyberattacks as a result of the rapid development and widespread usage of 5G, IoT, cloud computing, and other technologies. A network intrusion detection system (NIDS), the second line of defense behind the firewall, must correctly identify hostile network attacks, offer real-time monitoring and dynamic security measures, and devise tactics for countering them..
Information systems are now essential to any firm, regardless of its size or sector. The fact that information systems can store data and provide services, however, leaves them vulnerable to a variety of threats. Customized attacks on certain systems might have severe effects. In this climate, computer security research is becoming more and more important. High levels of security are required in modern living, and many strategies and practices have been developed to address these objectives. An intrusion detection system is one such instrument (IDS).

Intrusion detection systems (IDS) are meant to prevent a system from working normally by alerting administrators to suspicious activity on the network (IDS). Network-based, network-based intrusion detection systems (HIDS), and hybrid IDS are all subcategories of host-based intrusion detection systems (HIDS). Keeping a watch on network activity might help you catch any potentially harmful actions.

IDS systems are often built by activating promiscuous mode on the network interface card in order to record all network traffic. HIDS is used to keep track of encrypted data transfers to a certain host. It is powered by data obtained from a certain computer system. The traits of NIDS and HIDS are combined to form a composite IDS. Consequently, a network-based IDS can monitor the network and terminals.

James Anderson initially introduced the idea of intrusion detection in 1980, and machine learning methods were subsequently applied to find intrusions [1]. However, machine learning was disregarded at the time owing to restrictions on computer storage and processing capacity. Due to the fast expansion of computers, the introduction and promotion of artificial intelligence (AI), and other technologies, many academics are employing machine learning approaches to improve network security. [2-4] Some of these goals have been met.

II. RELATED WORK

James Anderson suggested the idea of incursion in 1980, and machine learning methods were used in the years that followed [1]. Machine learning was not taken into consideration at the time due to the storage and processing power of computers. Many academics have used machine learning approaches to enhance network security in light of the rapid growth of computers, the introduction and promotion of AI, and other technologies. They were able to accomplish a

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007014 |

few of their goals [2-4].
Support vector machines (SVM) were utilized by Parvez to examine the accuracy of classifier classification under To integrate feature selection and classification for the multi-class NSL-KDD Cup99 dataset in a new way, we looked at multiple dimensions of features [12]. [13] Shiraz explored and tested a variety of cutting-edge strategies for improving CANN's classification performance on the NSL-KDD Cup99 dataset.
KFN and KNN have the same data class rating, hence Second Nearest Neighbor (SNN) is used when they are the furthest and closest neighbors, respectively. The results show that CANN detection is successful and that failure rates are reduced while performance is maintained or improved. Using PCA and Firefly, Bhattacharya [14] proposes a machine learning technique.

The open data set collected from Cagle makes up the utilised data set. The IDS data set is first transformed using single-key encryption by the model, which is followed by dimensionality reduction using the PCA-Firefly hybrid method and classification using the XGBoost algorithm on the smaller data set.

Due to its ability to extract automated traits, deep learning has recently made significant improvements in the fields of computer vision, autonomous driving, and natural language processing (NLP) (NLP). For traffic categorization and intrusion detection, the application of deep learning is now under investigation. The task of detecting network traffic anomalies is transformed into a classification problem using deep learning and a training model [15]. Adaptive learning between normal network traffic and abnormal network traffic with extended training on sample data significantly improves real-time intrusion detection.
This made it possible for Torres et al. to change [16]. Additional use of Recurrent Neural Network (RNN) to categorize network traffic characteristics into a set of characters and its temporal elements to recognize phony network activity. A method for classifying malware The Wang Convolutional Neural Networks should be used to implement communications (CNN). Using the traffic's properties as pixels, a network traffic picture may be created. In order to classify the traffic, CNN then utilizes the picture.
The majority of studies balance the training set utilizing interpolation, oversampling, coding synthetic data, and other data augmentation approaches to enhance experimental performance. Even if their method effectively extends the minority class and mimics the data, the test data distribution may fall outside of the acceptable range. This distribution could not be correctly predicted by the classifier. We suggest using the DSSTE approach to increase or decrease the stable characteristics of the minority class, diminish the majority class in them, and delete important samples from an imbalanced training set. This technique reduces inconsistencies and produces reliable data.

III. PROPOSED METHODOLOGY

In order to balance out the training set and improve the intrusion detection system's classification accuracy The DSSTE technique is used to cope with network traffic that is uneven. Rather from increasing the number of majority samples, this approach actually decreases it. Random Forest, SVM, XGBoost, LSTM, Mini-VGGNet, and AlexNet are some of the classifiers we employ in our classification models.
Fig. 1 depicts our suggested intrusion detection model. Through the use of redundant, irregular, and missing value processing, our intrusion detection architecture exhibited data preprocessing. The training set should next be processed using the suggested DSSTE approach to balance the data after dividing the test and training sets. Prior to modeling, we use StandardScaler to digitize the sample labels and normalize the data. Following this, the classifier model is trained and tested using the altered training set.

A. MACHINELEARNINGANDDEEPLEARNINGALGORITHMS

A variety of methods, such as Random Forest and SVM as well as XGBoost and LSTM are used to build the classifier and test it in real-world data sets.
A long-term memory-based intrusion detection system was presented by Shamsinejad and Stoudemeyer [13]. In the KDD Cup99 dataset, you may discover information on DoS assaults and probing attacks. This model was developed by Kwan et al. [18], who have done a lot of research on deep learning models with an emphasis on data simplification (dimensionality reduction), classification, and other techniques [18].. It is shown that the FCN model is efficient in evaluating network traffic by contrasting it with conventional machine learning techniques. Using a composite feature selection strategy, Tama and colleagues [19] suggested a two-stage meta-classifier-based IDS system that produces correct feature representations.
Class imbalance has long been a challenge for machine learning. The identification of intrusions in highly varied groups of network traffic is thus very difficult. Due to this, a lot of researchers are trying to figure out how to identify asymmetric network traffic data invasions more accurately.

B. DSSTEALGORITHM

Different traffic data types in an imbalanced network have comparable representations, making it challenging for a classifier to discriminate between them during training. Particularly minorities are often the targets of attacks that

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007014 |

penetrate into daily life. Comparable examples show that repeating noise data is the most common feature of an uneven training set. Because the minority class is so large in the majority class, the classifier is unable to learn how it is distributed.

The characteristics are constantly changing, but the unique The traits that identify the minority group are consistent. In order to provide statistics that accurately represent the distribution, the stable traits of the minority class are boosted. For this reason, the DSSTE method is recommended for balancing the needs. Enneagrams may be altered by dividing training into groups based on their proximity to one other. As a result of the similarity of samples from neighboring groups,it is particularly challenging for a taxonomy to distinguish between classes. Because of this, we refer to cases in the groups of closest and furthest neighbors as tough and simple, respectively.

The challenging group's minority samples are then increased and decreased. In order to establish a new training group, minorities from the easy group and the challenging group are mixed with their reinforcement patterns. The scaling factor for the whole ENN method is based on K neighbors. A minority aggregation of strata collapses more rapidly than more complicated situations and more samples, in addition to the scaling factor K.

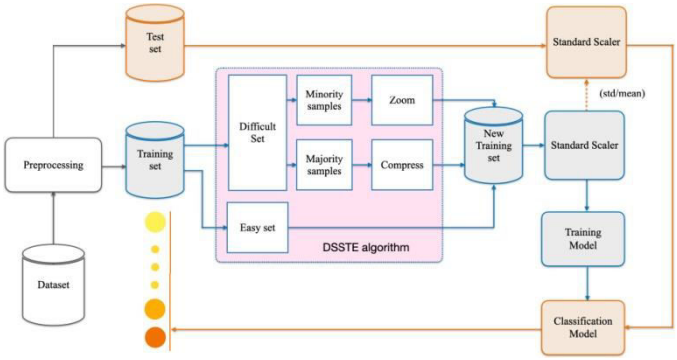
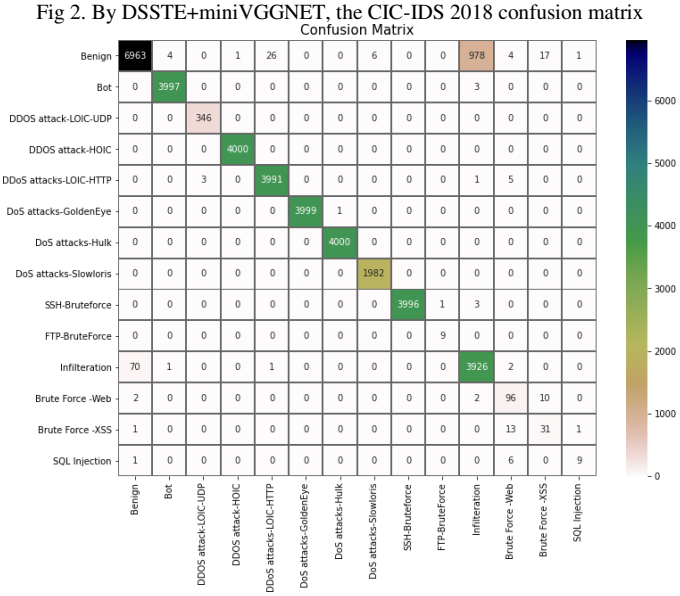


Fig 1.a model for network intrusion detection's general structure.

Fig 2. By DSSTE+miniVGGNET, the CIC-IDS 2018 confusion matrix



IV. IMPLEMENTAION

1. DATA PREPROCESSING

Due to extraction or input issues, some of the extracted dataset includes erroneous data, duplicate values, missing values, unbounded values, etc. So, we begin by processing data.

2. DATA TRANSFERRING

For a trained classifier to function, each record in the input data must be represented as a vector of real values. As a result, a numerical value is first given to each symbolic characteristic in the data collection. The KDD CUP 99 dataset,

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007014 |

for example, has both symbolic and numerical properties. TCP status flag, HTTP status, FTP status and ICMP status are all examples of protocol-specific symbolic attributes (for example, SF, REJ, etc.). Only numerical values for class property values are replaced by the method.

3. DATA NORMALIZATION

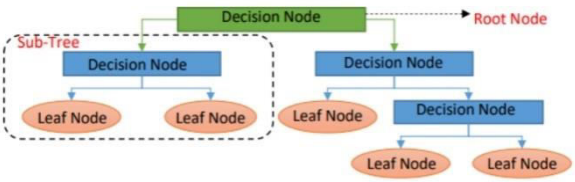
Normalization is a critical stage in data preparation that comes after all symbolic properties have been converted to numerical values. The technique of evaluating each feature's value within a proportionate range, known as data normalization, eliminates the bias in the data set toward characteristics with higher values. The information given in Section 5 is typical. The highest value of each attribute in the record is normalized, and they all lie within the same range [0-1]. This method also transfers and consolidates test data. To compare KDD Cup 99 with other systems evaluated against various sorts of assaults, we defined five categories. Unlike the other four classes, only one of them has standard records.

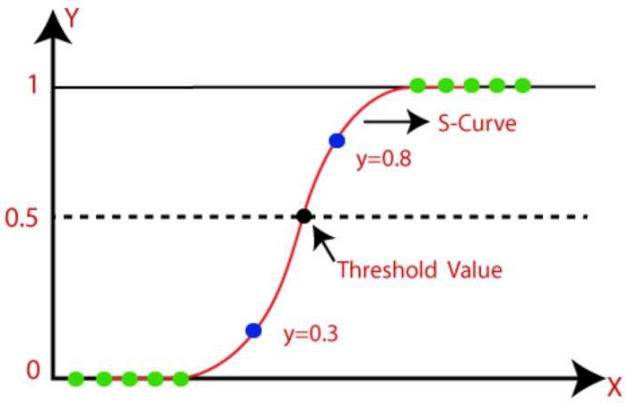
4. FEATURE SELECTION

Despite the fact that each connection in the data collection is represented by a unique set of properties, not all of this data is necessary to build an IDS. The most important aspects of the traffic data must be chosen for optimal performance. All that is required to address feature selection is Algorithm 1 from the prior section. It is possible to rank features based on their significance, but feature selection methods can't tell you how many features you'll need to train a classifier. When determining how many qualities are essential, this research follows the indicated method. Using the indicated feature selection method, this approach first ranks all features in order of importance.

5. DATA COLLECTION

a critical phase in the intrusion detection technique. When designing an intrusion detection system, it's important to consider the kind of data source used and where the data is collected. This research recommends a network-based IDS in order to verify our proposed methods and offer the best possible security for host or target networks. The recommended IDS examines the incoming network traffic on the router closest to the victim. It is based on this information that (s). All obtained data samples are first sorted according to transport/internet layer protocols and domain knowledge in order to prepare them for analysis in later stages. Protocol types are used to classify the data collected during the testing phase.





DECISION TREE

With each internal node representing a "test" on an attribute (for example, whether the coin comes up heads or tails),

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007014 |

the branches of this flowchart-like structure indicate the conclusion of that test, and the leaf nodes of this decision tree represent the class labels (a decision made after counting all the attributes). Classification criteria are shown using the path from the root to the leaf. Decision trees and impact diagrams are used in decision analysis as visual and analytical decision aids to determine the expected values (or expected utility) of competing alternatives. There are three types of nodes in a decision tree:

1. Random nodes, often represented by circles
2. Decision nodes, which often appear as squares
3. Final nodes, which are often shown as triangles

LOGISTIC REGRESSION IN MACHINE LEARNING

They may be yes or no, 0 or 1, true or false, and so forth. The only difference between logistic regression and linear regression is how they are used. To tackle classification and regression issues, linear and logistic regression are utilized. The regression line is replaced with a "S"-shaped logistic function in logistic regression (0 or 1). For example, whether or not the cells are cancerous or if the mouse is fat depending on its weight is shown by the logistic function curve. It is a popular machine learning technique because it can categorize fresh data using both continuous and discrete input. Logistic regression can rapidly find which characteristics perform best when categorizing observations using diverse data sources. The top image depicts the logistics function.

In the realm of supervised learning, logistic regression is one of the most often used machine learning algorithms. It is possible to forecast the categorical dependent variable by using a set of specified independent variables. Logistic regression is used to estimate the output of a categorical dependent variable. A discrete or categorical outcome is thus required. Probability values between 0 and 1 are returned instead of absolute values of 0 and 1.

Equation for Logistic Regression:

The logistic regression equation is derived from the linear regression formula. Logistic regression equations may be deduced using the following mathematical steps:

- We are aware that the equation for a straight line has the following form:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \cdots + b_nx_n$$

- In a logistic regression model, y can only be between 0 and 1, hence we may divide the following equation by (1-y):

$$\frac{y}{1-y}\text{; 0 for y= 0, and infinity for y=1}$$

- But because we require a range between -[infinity] and +[infinity], if we take the equation's logarithm, it becomes:

$$log\left[\frac{y}{1-y}\right] = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \cdots + b_nx_n$$

The final equation for logistic regression is the same as above.

Type of logistic regression:

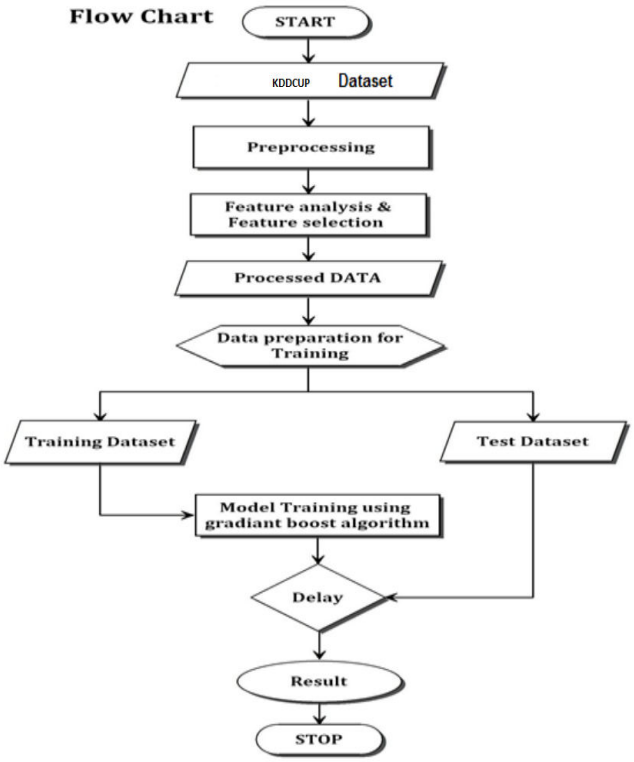
Three different forms of logistic regression can be distinguished based on categories:

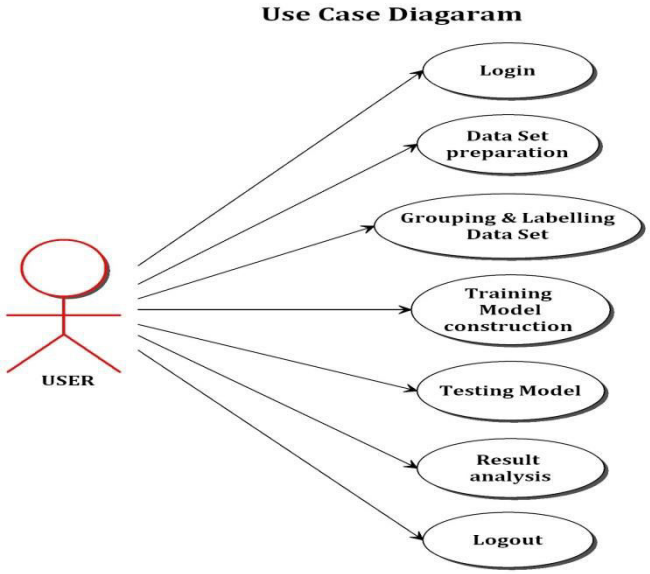
- **Binomial:**There are just two possible outcomes in binomial logistic regression: 0 or 1, pass or fail, and so on.
- **Ordinal:**There are just two possible outcomes in binomial logistic regression: 0 or 1, pass or fail, and so on.
- **Polynomial:**Any one of three or more unordered species, such as "cat," "dog," or "sheep," may serve as the dependent variable in polynomial logistic regression

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007014 |





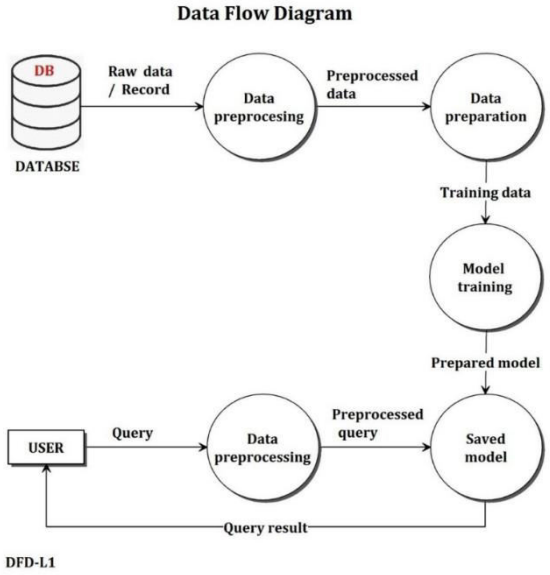
DATA FLOW DIAGRAM:

1. It is possible to visualize the flow of data in an information system using a data flow diagram (DFD). A data flow diagram is a visual representation of how data is processed in a computer system (structured design). It is common for designers to create a DFD at the context level that shows how the system interacts with external elements. DFDs show how data from external sources enters the system, travels between processes, and where it is logically stored. Only four symbols are used:
2. Boxes representing external entities that act as sources and destinations for data entry and exit.
3. In other methods, rounded rectangles representing processes may be labeled as "activities", "procedures", "procedures", "sub-systems", etc. These rounded rectangles accept, process, and output data as input.
4. Arrows depicting data flows, which may be electronic data that cannot be moved directly between physical objects or data stores; Instead, it must go through a process and external organizations are not allowed to directly access the data stores.
5. A three-sided flat rectangle represents data warehouses, which should be able to receive and deliver information for processing.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007014 |



SYSTEM TESTING

Because of this, testing is carried out. Testing is the process of finding errors or defects in the work. ensures that each component, subassembly, assembly, and/or finished product can be tested for proper operation. User expectations and demands must be met in order for a piece of software to be considered a success. Exams may be classified in a variety of ways. Each test kind responds in accordance with certain test requirements.

TYPES OF TESTS

Unit testing

By creating test cases for unit testing, you can make sure that the No matter what input is given to a computer program, it returns a correct answer. It is essential to evaluate the code's internal flow and all possible outcomes. It is the process of putting various pieces of application software through its paces. Before merging, this is done for each unit individually. It is a constant structural test that relies on the ability to comprehend its structure. When executing basic tests at the component level, unit tests examine a specific configuration of a system, application, or business process. In order to ensure that each step of a business process follows specifications and has well-defined inputs and outputs, unit tests are used.

Integration testing

Integration tests are performed on the merged software components to see whether they function as a single program. The exam is event-based and focuses more on a screen's or field's fundamental grade. Integration tests show that a group of components is correct and consistent even if individual components pass unit testing. Integration testing is designed particularly to draw attention to issues that occur while integrating components.

Functional test

Methodical assurance that the features being tested are available and fulfill all technical, commercial, and user manual requirements is provided by functional testing..
Focused on these areas is functional testing.
Valid input must be categorized in a manner that is acceptable.
Invalid Entry: The kinds of invalid entries listed above must be disregarded.
Use the indicated functions since they are required.
Output: Specific application output types must be checked.
Calling interacting systems or processes is necessary for systems and procedures.
Functional tests are developed and run with an emphasis on crucial requirements, functionalities, or unique test cases.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007014 |

Additionally, existing rules and follow-up procedures, business process flows, and systematic coverage of test data fields should all be taken into account. Before functional testing is finished, new tests are found that assess the value of the already-existing tests.

System Test

System testing ensures that the software system as a whole meets all of its specifications. He does a thorough assessment of the environment in order to ensure that he can obtain desired outcomes. System testing includes, for instance, configuration-based system integration testing. System testing focuses on pre-paid integration points and linkages and is based on process flows and descriptions.

White Box Testing

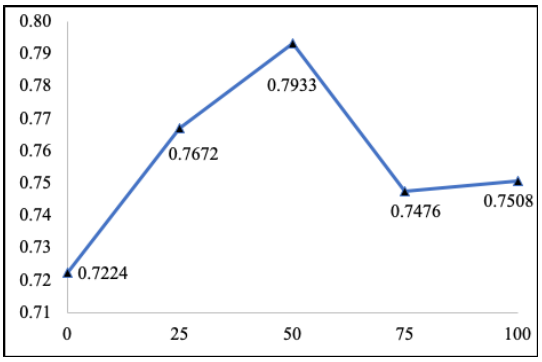
There are many ways to test software, but one of the most common is "white box testing," which refers to testing that is done by someone who already has a good understanding of how and why the program works. The goal is to get her. It's used to evaluate the parts of the black box level that aren't normally accessible.

Black Box Testing

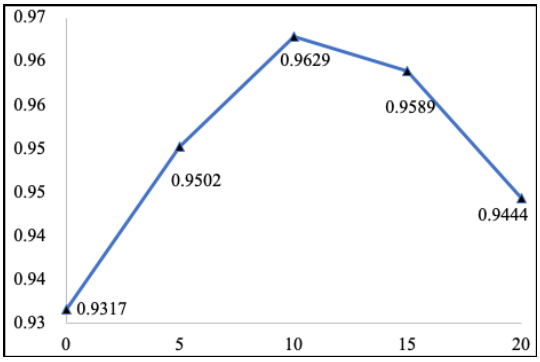
This kind of testing is known as "black box," since the testers don't know anything about the internal workings of the unit being tested. As with other types of testing, black box tests must be derived from a clear source document, such as a specification or requirements document. The application under test is treated as though it were a mysterious black box in this kind of testing. Is there a way to "look" inside the interior? A test creates input and interacts with output without regard to the program's functionality.

Unit Testing:

As part of the software development lifecycle, unit testing is often included into the coding phase, even if it isn't uncommon to separate the two tasks.



a)NSL-KDD KDDTest+



b)CSE-CIC-IDS2018

Test Environment Testing is part of the software development process. A testing procedure checks whether the product created meets the standards for which it was designed. Build test cases are one step in the testing process that includes

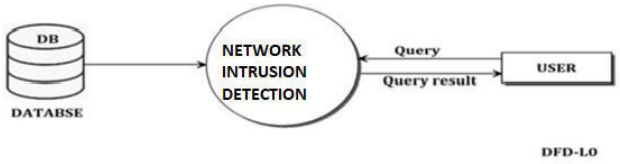
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007014 |

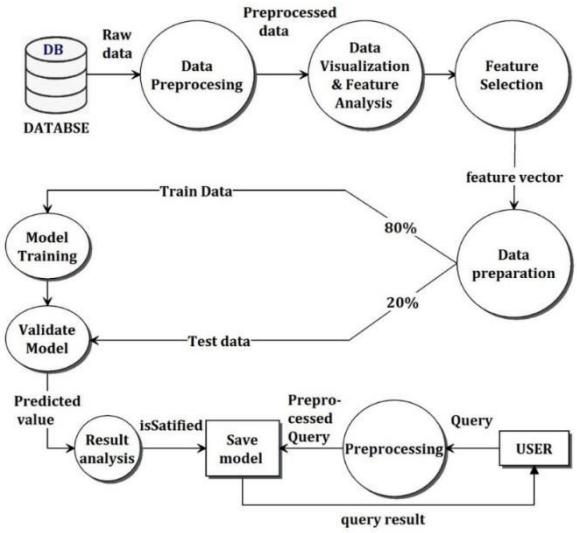
production testing. In our study, we want to use plant leaves to predict diseases. In addition, we classify the leaves according to the symptoms, causes and methods of treatment for each disease.

Unit Testing The core program logic is tested for accuracy and that the program's input generates valid output by designing unit test cases. Verifying internal code flow and all decision branches is crucial. Separate software components tailored to the application are tested after completion and before integration. It is a pervasive structural exam that depends on comprehension of its design. Unit tests examine a specific commercial application or system configuration while executing core tests at the component level. –> Using unit tests, you can make certain every step of the business process adheres to the specifications and explains the expected inputs and outputs in simple terms.



DFD-L0

Data Flow Diagram



DFD-L2

KNN(K NEAREST NEIGHBOUR)

- Step 1: How many neighbors are there? K.
- In the second stage, determine the Euclidean distance between K of your closest neighbors.
- Based on the Euclidean distance, choose K of your nearest neighbors.
- Next, count how many data points there are in each of the k categories.
- A new set of data is inserted into the category with the most neighbors as the last stage in the analysis.
- Step six: We've finished building our model.

V. CONCLUSION

By using the novel method for challenging group sampling (DSSTE) that was developed in this work, classification sampling may be enhanced. This method employs asymmetric network data learning. To improve minority learning under challenging patterns and decrease the imbalance in network traffic, a targeted increase in the number of minority patterns to be learnt may improve classification accuracy. We used a deep learning classifier with extra sampling techniques, together with six traditional machine learning methods. Studies demonstrate that our approach enhances attack detection more successfully by precisely sampling imbalanced network data that is propagating.

We discovered that deep learning outperforms machine learning in the experiment after sampling an unbalanced training set using the DSSTE method. Although neural networks boost data expressiveness and allow for automated feature extraction, existing public data sets already do so. Learning pre-processed features makes deep learning more challenging and makes automatic feature extraction less effective. By using actual network traffic data to train the deep

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 7, July 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1007014 |

learning model directly on feature extraction, we can preserve the advantages of deep learning while also reducing the effects of unbalanced data and improving classification accuracy. Foot

REFERENCES

1. D.E.Denning,''Anintrusion-detectionmodel,''*IEEETrans.Softw.Eng.*,vol. SE-13, no. 2, pp.222–232, Feb. 1987.
2. N. B. Amor, S. Benferhat, and Z. Elouedi, ''Naive Bayes vs decision treesinintrusiondetectionsystems,''in*Proc.ACMSymp.Appl.Comput.(SAC)*,2004, pp. 420–424.
3. M. Panda and M. R. Patra, ''Network intrusion detection using NaiveBayes,''*Int.J.Comput.Sci.Netw.Secur.*,vol.7,no.12,pp.258–263,2007.
4. M.A.M.Hasan,M.Nasser,B.Pal,andS.Ahmad,''Supportvec-tor machine and random forest modeling for intrusion detection system(IDS),'' *J. Intell. Learn.Syst. Appl.*, vol. 6,no. 1, pp.45–52, 2014.
5. N. Japkowicz, ''The class imbalance problem: Significance and strate-gies,''in *Proc. Int. Conf. Artif. Intell.*, vol. 56, 2000, pp. 111–117.
6. Y. LeCun, Y. Bengio, and G. Hinton, ''Deep learning,'' *Nature*, vol. 521,no. 7553, pp. 436–444, 2015.
7. Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, ''Deeplearning for visual understanding: A review,'' *Neurocomputing*, vol. 187,pp. 27–48, Apr. 2016.
8. T.Young,D.Hazarika,S.Poria,andE.Cambria,''Recenttrendsindeep learning based natural language processing [review article],'' *IEEEComput. Intell.Mag.*,vol.13,no. 3,pp.55–75,Aug.2018.
9. N.Shone,T.N.Ngoc,V.D.Phai,andQ.Shi,''Adeeplearningapproachtonetwork intrusion detection,'' *IEEE Trans. Emerg. Topics Comput. Intell.*,vol. 2, no. 1, pp. 41–50, Feb. 2018.

[1] D. A. Cieslak, N. V. Chawla, and A. Striegel, ''Combating imbalance innetwork intrusion datasets,'' in *Proc. IEEE Int. Conf. Granular Comput.*,May 2006, pp. 732–737.

[2] M. Zamani and M. Movahedi, ''Machine learning techniques for intru-sion detection,'' 2013, *arXiv:1312.2177*. [Online]. Available: http://arxiv.org/abs/1312.2177

[3] M. S. Pervez and D. M. Farid, ''Feature selection and intrusion classifi-cation in NSL-KDD cup 99 dataset employing SVMs,'' in *Proc. 8th Int.Conf.Softw.,Knowl.,Inf.Manage.Appl.(SKIMA)*,Dec.2014,pp.1–6.

[4] H. Shapoorifard and P. Shamsinejad, ''Intrusion detection using a novelhybrid method incorporating an improved KNN,'' *Int. J. Comput. Appl.*,vol. 173, no. 1, pp.5–9, Sep. 2017.

[5] S.Bhattacharya,P.K.R.Maddikunta,R.Kaluri,S.Singh,T.R.Gadekallu, M. Alazab, and U. Tariq, ''A novel PCA-firefly based XGBoost classifi-cationmodelforintrusiondetectioninnetworksusingGPU,''*Electronics*,vol. 9, no. 2, p. 219,Jan. 2020.

[6] A.Javaid,Q.Niyaz,W.Sun,andM.Alam,''Adeeplearningapproachfor network intrusion detection system,'' in *Proc. 9th EAI Int. Conf. Bio-inspiredInf.Commun.Technol.(FormerlyBIONETICS)*,2016,pp. 21–26.

[7] P.Torres,C.Catania,S.Garcia,andC.G.Garino,''Ananalysisofrecurrentneural networks for botnet detection behavior,'' in *Proc. IEEE BiennialCongr. Argentina (ARGENCON)*, Jun. 2016, pp.1–6.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.165

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details