Network Intrusion Detection System

[1] Sumathi. R
Department of CSE
Kalasalingam Academy of Research and Education
Anand Nagar, Krishnankoil, India
r.sumathi@klu.ac.in

[2] Paramesh. K
Department of CSE
Kalasalingam Academy of Research and Education
Anand Nagar, Krishnankoil, India.
9920004808@klu.ac.in

[3] Rohith. U
Department of CSE
Kalasalingam Academy of Research and Education
Anand Nagar, Krishnankoil, India
9920004138@klu.ac.in

[4] Anil Kumar.U
Department of CSE
Kalasalingam Academy of Research and Education
Anand Nagar, Krishnankoil, India
9920004263@klu.ac.in

[5] Lakshmi Chaitanya. P
Department of CSE
Kalasalingam Academy of Research and Education
Anand Nagar, Krishnankoil, India
9920004261@klu.ac.in

Abstract - Network intrusion discovery technology plays an important part in maintaining network security, the main work is to continuously descry the current network status, through the discovery of abnormal geste in the network state, timely warning to warn network directors. The punctuality and delicacy of the intrusion discovery system(IDS) is critical to the vacuity and trustability of the current network. In response to the problems of high false alarm rate, low discovery effectiveness and limited functions generally set up in IDS. With the growing reliance on networked systems and the adding complication of cyber pitfalls, the need for robust network intrusion discovery systems(NIDS) has come consummate. We're going to apply a Network Intrusion Discovery System that leverages both traditional machine learning and deep learning algorithms to enhance network security.collecting and preprocessing network traffic data, which serves as the foundation for training and testing the intrusion detection models. The objective of a project on a Network Intrusion Detection System (NIDS) using machine learning algorithms such as Decision Tree Classifier(DTC),K NearestNeighbours(KNN),LinearRegression(LR),MultiNomial NavieBayes(MNB),Random Forest Classifier(RFC),Support Vector Classifier(SVC) and deep learning algorithms such as convulutional neural networks(CNN),recurrent neural networks(RNN) and Feed forward neural networks(FNN)is to develop a system that can effectively detect the attack type. Use machine learning and deep learning algorithms to identify abnormal patterns and behaviors in network traffic dataset.Identification of Intrusions: Detect and classify different types of network intrusions

Keywords—Decision Tree Classifier,Linear Regression,K NearestNeighbours,MultiNomial NavieBayes,Random Forest Classifier,Support Vector Machines,Convulutional neural networks,recurrent neural networks,feed forward neural networks

INTRODUCTION

A network intrusion detection system (NIDS) is a pivotal security mechanism designed to monitor and analyze network traffic for potential threats and unauthorized access attempts. Its primary objective is to detect and respond to suspicious activities, malicious behaviors, or security breaches within a computer network.NIDS operates by examining inbound and outbound traffic, scrutinizing packets, and identifying patterns or anomalies that deviate from established network norms or predefined signatures.
It employs various detection methods, including signature-based detection, which relies on a database of known attack patterns, and anomaly-based detection, which identifies deviations from expected behavior.[3-5]

The system plays a crucial role in safeguarding networks by providing real-time alerts or initiating automated responses upon detecting potentially harmful activities. It enhances overall network security by swiftly identifying and mitigating threats, thus preventing potential damage or unauthorized access to critical systems and data. A Network Intrusion Detection System (NIDS) fortified with Machine Learning (ML) and Deep Learning (DL) algorithms represents a cutting-edge approach to fortifying network security. This project amalgamates the power of advanced computational techniques with the intricate analysis of network traffic patterns to proactively detect and mitigate potential cyber threats.Unlike traditional NIDS reliant on predefined signatures or rule-based systems, ML and DL-empowered NIDS leverage sophisticated algorithms to autonomously learn and adapt to evolving cyber threats. These algorithms can identify anomalies, detect sophisticated attack patterns, and distinguish between normal and malicious network activities by analyzing vast amounts of network data.[7-8]

LITERATURE SURVEY

Several studies have significantly contributed to the field of network intrusion detection. One survey offers an extensive overview of intrusion detection techniques, emphasizing their respective strengths and weaknesses. Another study focuses on anomaly-based detection using autoencoders, showcasing their effectiveness in identifying network anomalies. Furthermore, deep learning methods, including CNNs and LSTMs, have been extensively reviewed for their potential in handling complex attacks. Comparative studies have evaluated the accuracy and computational aspects of various machine learning algorithms for intrusion detection. Additionally, specific network architectures like LSTM networks and stacked autoencoders have been explored for capturing temporal patterns and feature extraction. Lastly, a survey outlines the evolutionary shift towards machine learning and deep learning methods in intrusion detection systems.[3] In recent advancements within Network Intrusion Detection Systems (NIDSs), a multi-stage optimized framework has been proposed to reduce computational needs while maintaining tracking efficacy

ML-based structure significantly diminishes training sample dimensions (by 74%) and feature set sizes (by 50%) using datasets like CICIDS 2017 and UNSW-NB 2015.Another avenue, highlighted in , emphasizes the limitations of existing datasets in adapting to evolving threats. This study

underscores the absence of genuine network threats in recent datasets, hindering the effectiveness of machine learning-based IDS techniques. The paper advocates for a categorization and dataset analysis approach to enhance current research landscapes and combat sophisticated threats, acknowledging the necessity for dataset improvements to align with modern attack patterns and bolster system defenses.[2]

The trajectory of NIDS evolution has been multifaceted. Signature-based systems, while effective against known threats, grapple with the agility of zero-day attacks. Anomaly-based systems initially promised adaptability but were plagued by high false positives. The convergence toward hybrid systems leverages the best of both worlds, employing signature databases alongside machine learning algorithms to dynamically adapt to new attack vectors. Signature-based NIDS operate on a blacklist approach, efficiently detecting known patterns but faltering against previously unseen threats. Anomaly-based systems, employing statistical models, identify deviations from normal behavior but often struggle to define 'normalcy' accurately. Hybrid models, integrating machine learning techniques like clustering algorithms and neural networks, excel in detecting evolving threats while reducing false positives through adaptive learning. [5]

Behavioral analysis in NIDS involves monitoring network activities to establish baselines of normal behavior and subsequently identify deviations that may indicate malicious activity. This method, while effective in detecting novel attacks, faces challenges in defining what constitutes "normal" in dynamic and diverse network environments. Recent research explores dynamic behavioral profiling and machine learning algorithms to continuously adapt to evolving network behaviors, improving accuracy and reducing false positives.Machine Learning Techniques for Anomaly Detection: Machine learning techniques play a pivotal role in anomaly detection within NIDS. Traditional approaches have relied on statistical methods, but recent advancements in machine learning, such as unsupervised learning algorithms like clustering and autoencoders, have shown promise in detecting anomalies without predefined labels. These techniques leverage the inherent patterns in network data to identify abnormalities, thus improving the detection of previously unseen threats.[1]

METHODOLOGY

Network Intrusion Detection System website is developed front end is developed using css,html,javascript and the back end is developed using flask and python.Users Data Base will be maintened by stroing the information of the user The website loads with the index page showing the Login and SignUp buttons and when the user signup and login then the home page will be opened where UPLOAD menu,NSL_KDD menu,ML_MODELS menu,DL_MODELS menus are present.The machine learing models such as Decision Tree Classifier,LinearRegression,KNearestNeighbours,MultiNomial NavieBayes,RandomForest Classifier,Support Vector Machines,and deep learning models such as Convulutional neural networks,recurrent neural networks,feed forward neural networks are developed and deployed into the system to classify the attacktype.Our proposed system uses the Hard Voting Method and the final result is the atatck of a which had the highest probability of being predicted by each of the classifiers.Among all the predicted attacks the attack which has the highest frequency is taken as the final result if no attack is common among the predicted attacks then the attack which is predicted by the algorithm which gives the higher accuracy is taken as the final result.the data set used to train and test the models is NSL-KDD data set.

Table 1 Current Parameter values taken for data set

Attribute Number	Attribute
1	Dst_host_srv_serror_rate
2	Service
3	Flag
4	Dst_host_serror_rate
5	Diff_srv_rate
6	Dst_host_same_src_port_rate
7	Serror_rate
8	Wrong_fragment
9	Protocol_type
10	Logged_in
11	Srv_serror_rate
12	Dst_host_same_srv_rate
13	Is_host_login
14	Is_guest_login
15	Srv_diff_host_rate
16	Dst_host_srv_count
17	Count
18	Same_srv_rate
19	Dst_host_srv_diff_host_rate
20	Srv_count
21	Dst_host_diff_srv_rate
22	Num_failed_logins
23	land

NSL_KDD data is cleaned and preprocessed and the above important features are obtained by combining the feature selection techniques such as recursive feature elimination,forward selection and backward elimination and correlation based feature selection.After preprocessing of the data set the machine learning and deep learning models are trained and tested and their accuracies are obtained by test data set.All the ML and DL models are then deployed into the website to make predictions about the attack type in the network.

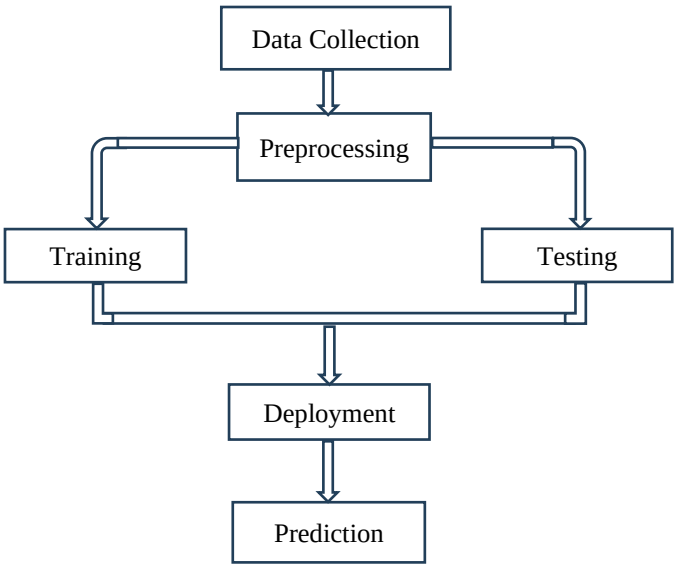

Fig 1. System Design

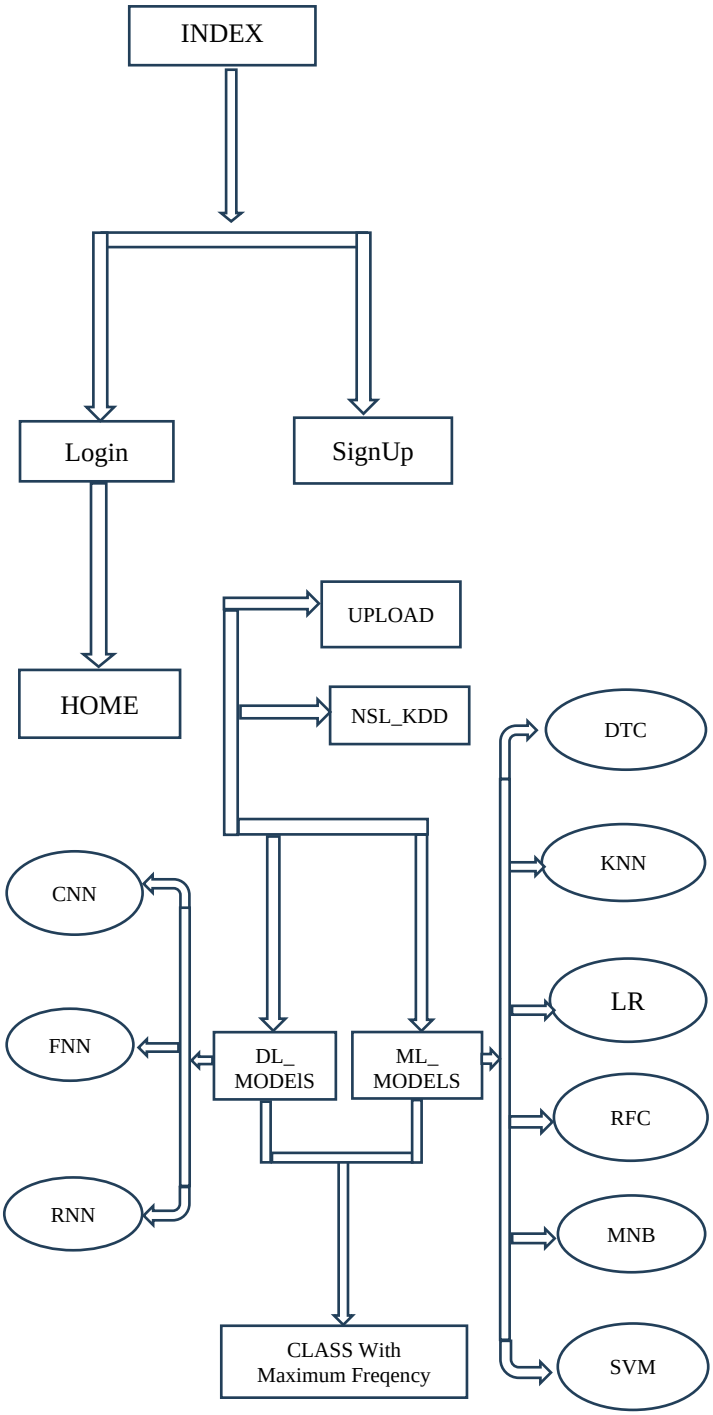

Figure 2. Proposed Flow Chart

When the user opens the website the index page will be vissible at the top right corner login and signup buttons are present the user needs to create account by signup by verifying the otp which will be sent to the user given gmail while signup and then otp verification will be done and the user account is created with the user name and password .The user have to login into the page and when the user logins then it will open the home page where the user can see the menus UPLOAD,NSL_KDD,ML_MODELS,DL_MODELS where the user can perform several operations by clicking on the menus.The modules/functionalites created in this project are SIGNUP,LOGIN,UPLOAD,NSL_KDD,ML_MODELS,DL_MODELS.Each Module can perform specific task.Users data is stored at the backend in a database the data base is created using flask and python the front end is created using html,css and javascript.at the back end user information can be accessed by writing query by using sql programming language as the sql language can work fastly and in an efficiant manner.

A.SIGNUP

In signup page the details of the user such as name,email,phonenumber,username,password are obtained and an otp is sent to given email when the user enters the correct otp and click on the button signup then only the account will be created other it will display invalid otp.

B.LOGIN

The user can login into the home page by entering the username and password.when the user login successfully then the home page will be opened and at the top welcome to home page this is the protected home page that you can access will be scrolling from right to left

C.UPLAOD

When the user clicks the upload menu choose file and upload data set buttons displays the choose file button is clicked and then select a csv file and then click on the upload button then the entire analsis report of the data set will be displayed in an iframe

D.NSL_KDD

It consists of the features and their description of the data set where we are able to search the feature and can see the infromation about the specific feature.

E.ML_MODELS

It consists of different features.the input values for the features is given and then we have to click on the select the machine learning algorithm button where different machine learning algorithms are used such as Decision Tree Classifier(DTC),KNearestNeighbours(KNN),LinearRegression(LR),MultiNomialNavieBayes(MNB),RandomForestClassifier(RFC),Support Vector Classifier(SVC) and then click on the button predict attack type.so the values given by the user will be taken as the input of the models and at the back end the machine learing algoirthms processes the input data and then gives the predicted attack type there are various attack types are present in the network.

F.DL_MODELS

It consists of different features.the input values for the features is given and then we have to click on the select the deep learning algorithm button where different deep learning algorithms are used such as such as convulutional neural networks(CNN),recurrent neural networks(RNN) and Feed forward neural networks(FNN) and then click on the button predict attack type.so the values given by the user will be taken as the input of the models and at the back end the deep learing algoirthms processes the input data and then gives the predicted attacktype.The attack different attack types it may predict are apache2, back, buffer_overflow, ftp_write, guess_passwd,httptunnel, imap, ipsweep, land, loadmodule, mailbomb,mscan, multihop, named, neptune, nmap, normal, perl, phf, pod, portsweep, processtable, ps, rootkit, saint,satan, sendmail, smurf, snmpgetattack, snmpguess, spy, sqlattack, teardrop, udpstorm, warezclient, warezmaster, worm, xlock, xsnoop, xterm.each algorithm processes the data and give result which may be same or different.

RESULTS AND DISCUSSIONS

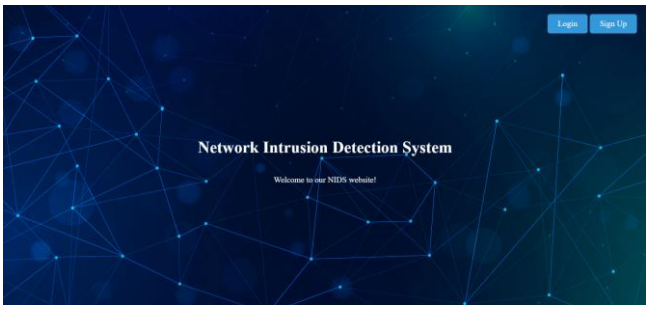

Figure 3. index page

Index page consists of the text Network Intrusion Detection System and welcome to our NIDS website.It consists of SignUp and Login buttons where the user can create account and login respectively.

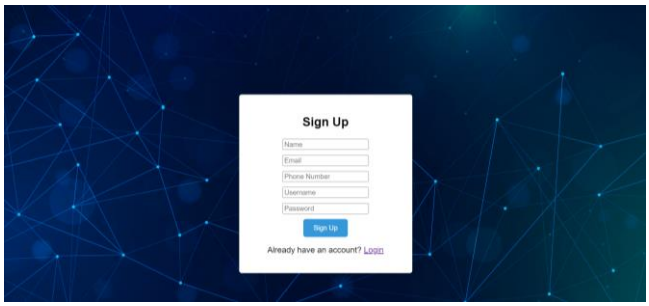

Figure 4. Sign Up page

The SignUp Page will get the information of the user and creates the account to the user.

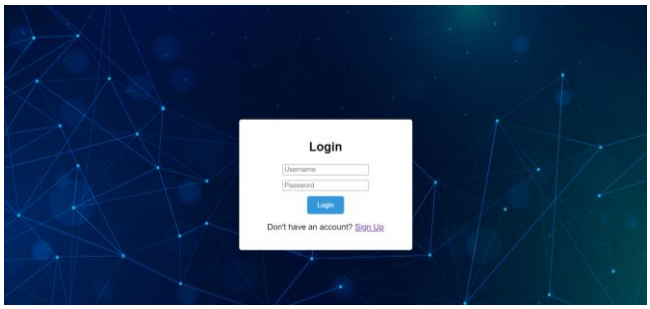

Figure 5. Login page

The user can enter the username and password and login into the home page

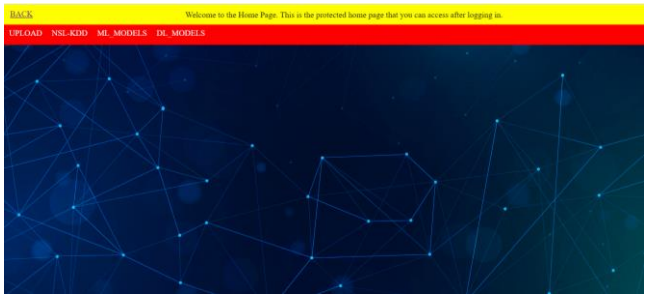

Figure 6. Home Page

In home page UPLOAD,NSL-KDD,ML_MODELS,DL_MODELS menus are present

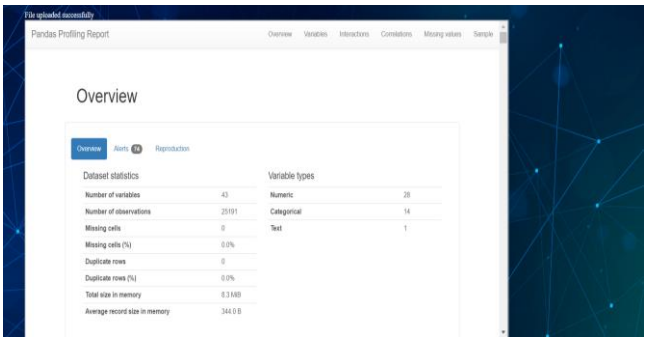

Figure 7. Analysis Report

Analysis report is generated using pandas profiling library when the user successfully upload the csv file and click on the upload button then the data set is processed at the back end and the pandas profiling report is generated and it will be displayed inside the I frame.

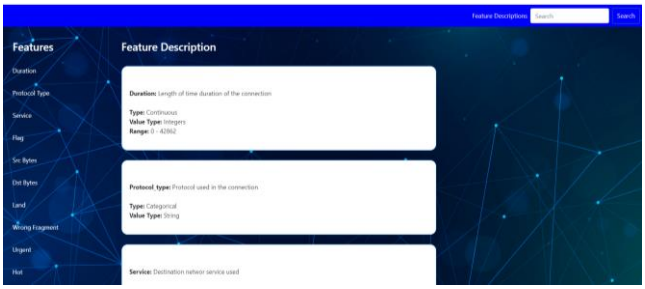

Figure 8.dataset description

We can search the specific feature by click on the search button and know the information about the specific features

Figure 9.feature values

Figure 10.remaining feature values

User will have to enter the input data used for prediction of attack.

Figure 11.machine learning algorithms

machine learning algorithms are used such as Decision TreeClassifier(DTC),KNearestNeighbours(KNN),Linear Regression(LR),MultiNomialNavieBayes(MNB),RandomForestClassifier(RFC),Support Vector Classifier(SVC) are developed and deployed into the system to predict the attack type.

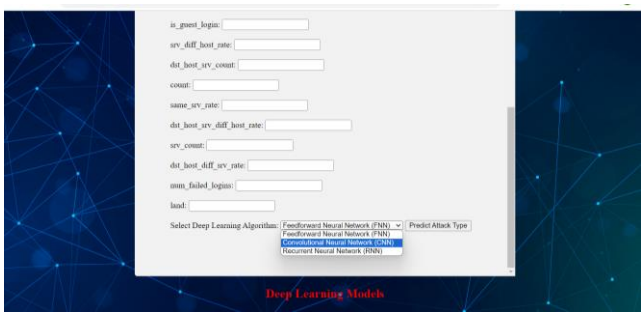

Figure 12.deep learning algorithms

deep learning models such as Convulutional neural networks,recurrent neural networks,feed forward neural networks are developed and deployed into the system to classify the attacktype.

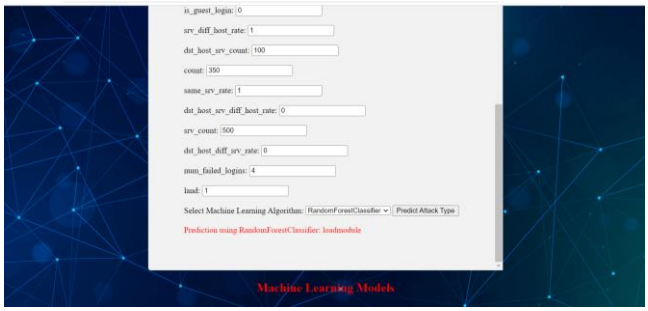

Figure 13.prediction using ml algorithms

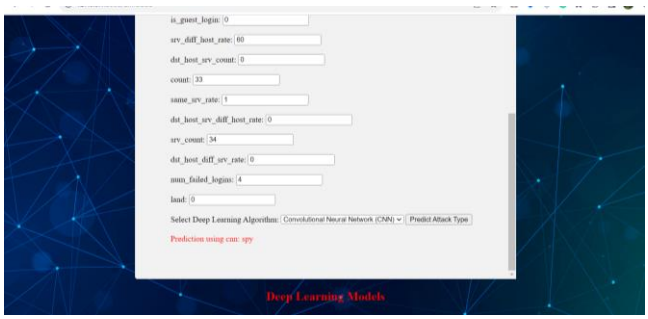

Figure 14.prediction using dl algorithms

Table 2 algorithms and their accuracies

Algorithm	Accuracy
1.RandomForestClassifier	98.1988
2.KNeighboursClassifier	97.8218
3.DecisionTreeClassifier	97.8016
4.FeedForwardNeuralNetworks	97.5200
5.ConvulutionalNeuralNetworks	97.0000
6.SupportVectorMachines	96.0173
7.LogisticRegression	95.2194
8.MultinomialNavieBayes	90.8598

The above accuracies are obtained for each of the algorithms by training and testing the model with the NSL-KDD data set.randomforestclassifier has got the highest accuracy.The user will give the input values for the features and the algorithms process the data.the final predicted attack type is the attack type of which majority of the algoirthms predicted.the among the predicted attack types by the algorithms the attack type whose frequency is maximum is considered as the final prediction result.if there is no common attack among the predicted attacks by the above algorithms then attack type predicted by the random forest classifier is taken as the final result.

CONCLUSIONS

network intrusion detection using machine learning and deep learning algorithms offers promising solutions to bolster cybersecurity measures. Through the utilization of various algorithms, such as decision trees, random forests, support vector machines, and neural networks, this project aims to enhance the accuracy and efficiency of detecting anomalous activities within networks.By leveraging machine learning models, the system can learn patterns from network data and distinguish between normal and malicious behavior, thereby fortifying preemptive measures against potential cyber threats. Deep learning algorithms, particularly neural networks, exhibit the capacity to analyze intricate network data, uncover hidden patterns, and adapt to evolving threats in real-time.This project's potential lies in its ability to continuously improve its detection capabilities through iterative learning, enabling the system to stay updated and resilient against emerging intrusion techniques. Furthermore, the integration of machine learning and deep learning techniques allows for a more comprehensive and sophisticated approach to identifying various types of network intrusions, including intrusion attempts, malware attacks, and unauthorized access.In conclusion, the integration of machine learning and deep learning algorithms in a network intrusion detection system presents a powerful framework to proactively safeguard networks against cyber threats. Its success relies on continuous refinement, adaptability, and a proactive approach to address the evolving landscape of cybersecurity threats.

REFERENCES

[1]R. Lippmann, J. Haines, D. Fried, J. Korba and K. Das. ”The 1999 DARPA off-line intrusion detection evaluation”. Computer networks, vol. 34, no. 4, pp. 579 595, 2000.

[2] W. Lee and S. Stolfo. ”A framework for constructing features and models for intrusion detectionsystems”. ACM transactions on information and system security, vol. 3, no. 4, pp.227261,2000.DOIhttp://dx.doi.Org/10.1145/382912.3829

[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] M. Vladimir, V. Alexei and S. Ivan. ”The MP13 approach to the KDD’99 classifier learning contest”. SIGKDD explorations newsletter, vol. 1, pp.
76 77, 2000. DOI http://dx.doi.org/10.1145/846183. 846202.
[5] R. Agarwal and M. Joshi. ”PNrule: A new framework for learning classier models in data mining”. Tech. Rep. 00-015, Department of Computer Science, University of Minnesota,.
[6]. G. Karatas, O. Demir and O. K. Sahingoz, "Increasing the Performance of Machine Learning-Based IDSs on an Imbalanced and Up-to-Date Dataset," in IEEE Access, vol. 8, pp. 32150-32162, 2020, doi: 10.1109/ACCESS.2020.2973219.
[7]. Njogu, H. W., & Jiawei, L. (2010, July). Using alert clusters to reduce IDS alerts. In 2010
3rd International Conference on Computer Science and Information Technology (Vol. 5, pp. 467-471). IEEE.
[8]. R. K. Vigneswaran, R. Vinayakumar, K. P. Soman and P. Poornachandran, "Evaluating Shallow and Deep Neural Networks for Network Intrusion Detection Systems in Cyber Security," 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Bengaluru, India, 2018, pp. 1-6, doi: 10.1109/ICCCNT.2018.8494096.
[9]. I. Ahmad, M. Basheri, M. J. Iqbal and A. Rahim, "Performance Comparison of Support
Vector Machine, Random Forest, and Extreme Learning Machine for Intrusion Detection,"
in IEEE Access, vol. 6, pp. 33789-33795, 2018, doi: 10.1109/ACCESS.2018.2841987. J. P. Anderson *et al.*, “Computer security threat monitoring and surveillance,” Technical report, James P. Anderson Company, Fort Washington, Pennsylvania, Tech. Rep., 1980.
[10] T. Vaidya, “2001-2013: Survey and analysis of major cyberattacks,” *arXiv preprint arXiv:1507.06673*, 2015.
[11] W. Lee, W. Fan, M. Miller, S. J. Stolfo, and E. Zadok, “Toward costsensitive modeling for intrusion detection and response,” *Journal of computer security*, vol. 10, no. 1-2, pp. 5–22, 2002.
[12] Z. C. Lipton, J. Berkowitz, and C. Elkan, “A critical review of recurrent neural networks for sequence learning,” *arXiv preprint arXiv:1506.00019*, 2015.
[13] E. Hodo, X. Bellekens, A. Hamilton, C. Tachtatzis, and R. Atkinson, “Shallow and deep networks intrusion detection system: A taxonomy and survey,” *arXiv preprint arXiv:1701.02145*, 2017.
[14] N. Gao, L. Gao, Q. Gao, and H. Wang, “An intrusion detection model based on deep belief networks,” in *Advanced Cloud and Big Data (CBD), 2014 Second International Conference on*. IEEE, 2014, pp. 247–252.
[15] M. Moradi and M. Zulkernine, “A neural network based system for intrusion detection and classification of attacks,” in *Proceedings of the IEEE International Conference on Advances in Intelligent SystemsTheory and Applications*, 2004, pp. 15–18.
[16] S. Mukkamala, A. H. Sung, and A. Abraham, “Intrusion detection using an ensemble of intelligent paradigms,” *Journal of network and computer applications*, vol. 28, no. 2, pp. 167–182, 2005.
[17] J.-S. Xue, J.-Z. Sun, and X. Zhang, “Recurrent network in network intrusion detection system,” in *Machine Learning and Cybernetics, 2004. Proceedings of 2004 International Conference on*, vol. 5. IEEE, 2004,
pp. 2676–2679.