IJIRCCE

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

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 5, May 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205078 |

Network Intrusion Detection System Using Machine Learning Techniques

Pratik Tonge, Prof Neehal Jiwane, Prof Lowlesh Yadav

Department of Computer Science and Engineering, Shri Sai College of Engineering and Technology, Chandrapur, India

Department of Computer Science and Engineering, Shri Sai College of Engineering and Technology, Chandrapur, India

Department of Computer Science and Engineering, Shri Sai College of Engineering and Technology, Chandrapur, India

ABSTRACT: The number of attacks over the Internet has increased over the years due to the advancement and easy availability of computing technologies. Attackers discover new attack types therefore to prevent these attacks firstly they need to be identified correctly by intrusion detection systems (IDS). An intrusion Detection System (IDS) is needed to make the network secure. The supervised machine learning system is designed to scan network traffic whether it is malicious or benign. To have best intrusion attack detection success rate, a combination machine learning algorithm and feature selection method has been used. In this paper, we have used the machine learning algorithm Artificial Neural Network (ANN) with feature selection from network traffic. We have used NSL- KDD dataset to classify network traffic using ANN supervised machine learning techniques.

KEYWORDS: IDS, IoT, Anaconda, Python, SQL,Spyder,Wireshark

I. INTRODUCTION

Intrusion Detection System is a software application to detect network intrusion using various machine learning algorithms.IDS monitors a network or system for malicious activity and protects a computer network from unauthorized access from users, including perhaps insider. The intrusion detector learning task is to build a predictive model (i.e. a classifier) capable of distinguishing between 'bad connections' (intrusion/attacks) and a 'good (normal) connections'. With the outstanding boom of the web technologies, numerous forms of cybercrimes are growing rapidly. As example, stealing of expertise, phishing, carding, viruses, economic fraud, intrusions attacks. Network assaults is one among crime kinds that intends to compromise the confidentiality,integrity and accessibility of the data in organization computer traffic or in the local host. Intrusion detection system plays an important role in protecting organization systems from different cyberattacks. IDS works on network of the system and act as defense mechanism to secure the organization computer systems.The intrusion detection system may be host-based IDS(HIDS) or network-based IDS (NIDS). The host-based intrusion detection systems are adopted by network administrators to monitor and analyze activities on a particular machine. New Technologies like AI and ML are also evolving and provide great results. There are various algorithms and techniques in these technologies which can work on large data sets and produce optimum results. Intrusion detection systems combined with this ML algorithms can monitor and analyze computer systems large network traffic easily and efficiently detect intrusion attacks.The scope of this project is to, it may help to many people to use better network identification method which is an important for network traffic monitoring, data analysis and is the key to improve the quality of user service. This project also shows that proposed method is get the best performance and improve the intrusion detection as compare to current methods. The device can decide whether or not a code or a document is malicious or now no longer in a completely small quantity of time without the want of separating it in a sandbox to carry out the analysis. It is likewise beneficial for the detection of more and more polymorphic nature of malware. It causes noticeably correct category of malicious web pages and additionally the high fake effective rate. This device can not only detect changed or variation present malware however also can detect the preceding unknown attack.

The goal of a network intrusion detection system is to discover unauthorized access to a computer network by analyzing traffic on the network for signs of malicious activity.An intrusion detection and prevention system (IDPS) is a solution that monitors a network for threats and then takes action to stop any threats that are detected.

The internet is the hub for information exchange which transports trillions of bytes a day. Attackers continuously develop new exploits and attack techniques designed to circumvent your defenses. No firewall is full proof, and no

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205078 |

network is impenetrable, attackers find new ways to perform intrusion attack and get into system. A network intrusion detection system is crucial for network security because it enables you to detect and respond to malicious traffic. Intrusion detection systems have become a necessary component of almost every network security infrastructure now-a-days. So far, several IDS's have been proposed and they all have their own limitations. In existing system no any machine learning algorithm are used for intrusion detection. Using ML algorithms in the intrusion detection system can help to improve the accuracy and efficiency of detecting intrusion attack. So, a smart intrusion detection system is needed to detect attacks efficiently so computer system remains safe and secure.

II. LITERATURE SURVEY

In today's world of network communications, network infiltration is the most pressing worry. Network assaults are becoming more common, posing a serious threat to network services. Various studies have already been carried out in order to establish an effective and efficient method to prevent network intrusion and protect network security and privacy. A mixture of two machine learning methods is suggested in this study to classify any unusual behavior in network traffic. The detection accuracy, false positive rate, false negative rate, and time taken to detect the incursion are all evaluated to determine the overall efficiency of the proposed technique. The suggested technique shows that the algorithm is successful in identifying intrusions with improved detection accuracy.[2]

Various studies have already been carried out in order to establish an effective and efficient method to prevent network intrusion and protect network security and privacy. A mixture of two machine learning methods is suggested in this study to classify any unusual behavior in network traffic. The detection accuracy, false positive rate, false negative rate, and time taken to detect the incursion are all evaluated to determine the overall efficiency of the proposed technique.[3]

Secure automated threat identification and prevention is a more effective method of reducing analyst effort by analyzing the network and server operations and alerting the analyst if any suspicious behavior is discovered in network traffic. It continually monitors the system and responds in accordance with the threat environment. The reaction action differs depending on the phase.Suspicious activities are discovered using artificial intelligence that operates as a virtual analyst inconjunction with a network intrusion detection system to defend against the threat environment andtake relevant steps with the analyst's consent. In the last stage, packet analysis is used to search forattack vectors and then classify supervised and unsupervised data.[1]

A comprehensive survey of some major techniques of machine learning implemented on intrusion Detection was done where techniques based on K-means, K-means with principal component analysis, Random Forest algorithm extreme learning the machine, techniques, classification algorithms such as Naive Bayes algorithm, Hoeffding Tree Algorithm. Also, Accuracy Updated Ensemble algorithm, Accuracy Weighted Ensemble algorithm, Support Vector Machine, Genetic algorithm, and Deep learning were studied.[5].

III. METHODOLOGY

The system we have created is to detect the intrusion attack by scanning the network traffic of the system. The system has feature selection and learning algorithm. Feature selection component are responsible to extract most relevant features or attributes to spot the instance to a specific group or class. The machine learning algorithm component builds the required intelligence or knowledge using the result found from the feature selection component. The machine learning model is trained on dataset and it builds its intelligence. Then the learned intelligences are applied to the testing dataset to live the accuracy of home much the model correctly classified on unseen data. We have used ANN algorithm and it is trained on NSL- KDD dataset. We use Wireshark to track the network packets.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205078 |

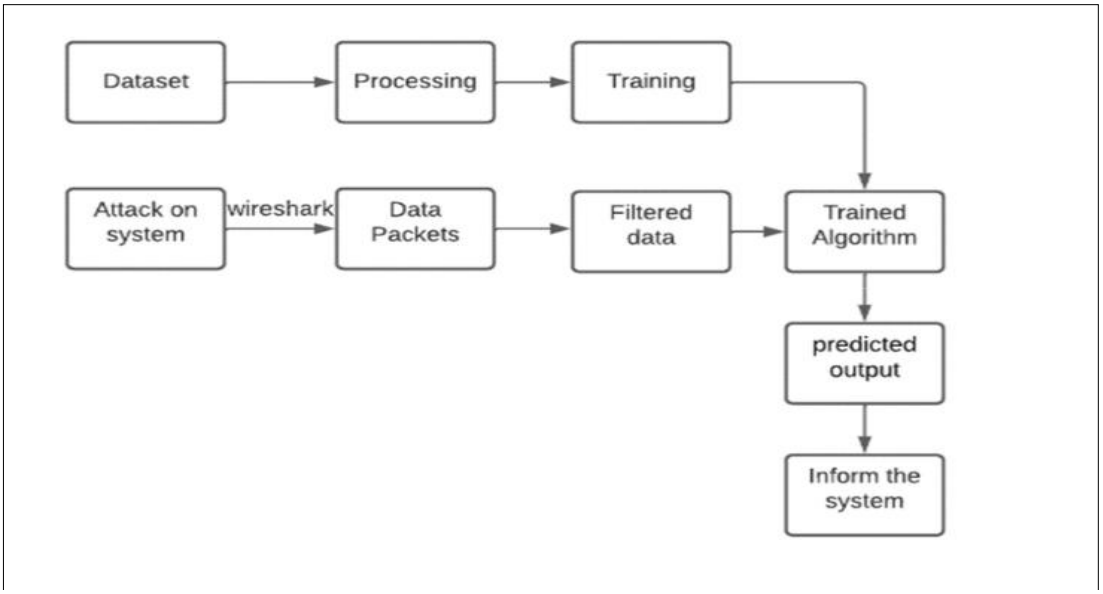

Fig No.1 Block Diagram of Mobile Cloud Computing

Flowchart

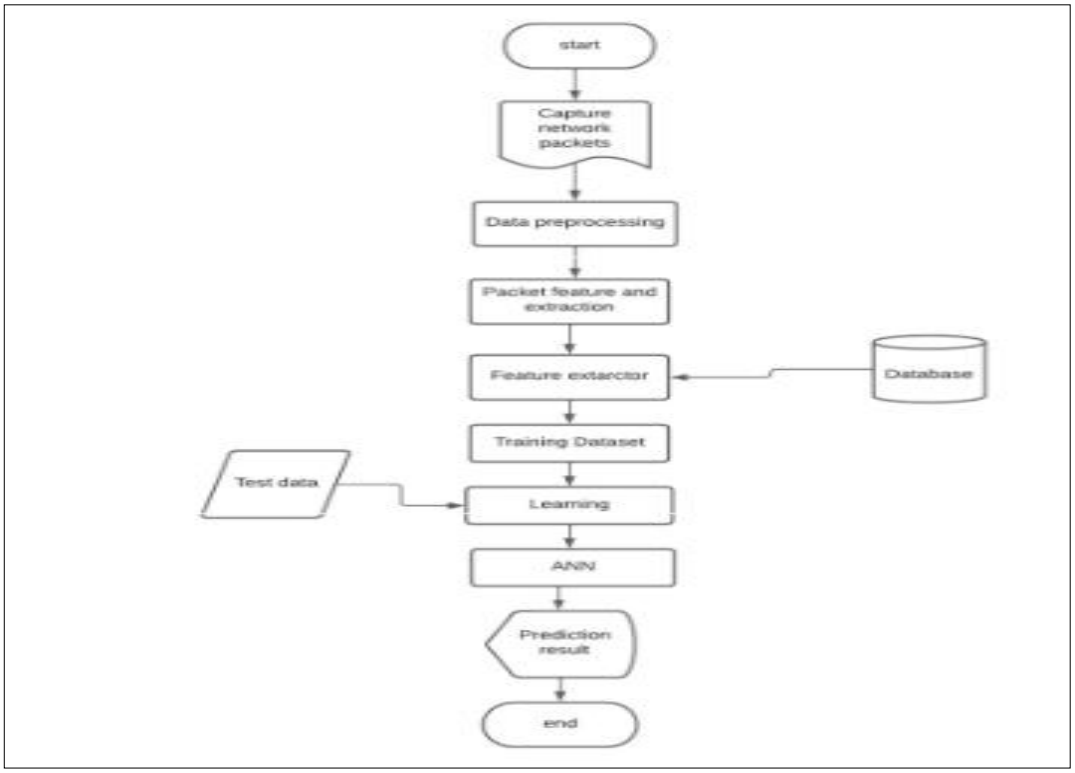

Fig No.2Flowchart of Mobile Cloud Computing

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205078 |

IV. WORKING

The system we have created is to detect the intrusion attack by scanning the network traffic of the system. The system has feature selection and learning algorithm. The machine learning model is trained on dataset and it builds its intelligence. We have used ANN algorithm and it is trained on NSL- KDD dataset. We have created a local host website on which we perform attacks while the attack is being executed Wireshark is used to capture network packets these network packets are filtered and given to the machine learning algorithm ANN. ANN then analyzes the filtered packets and predicts the type of attack happened on the system.

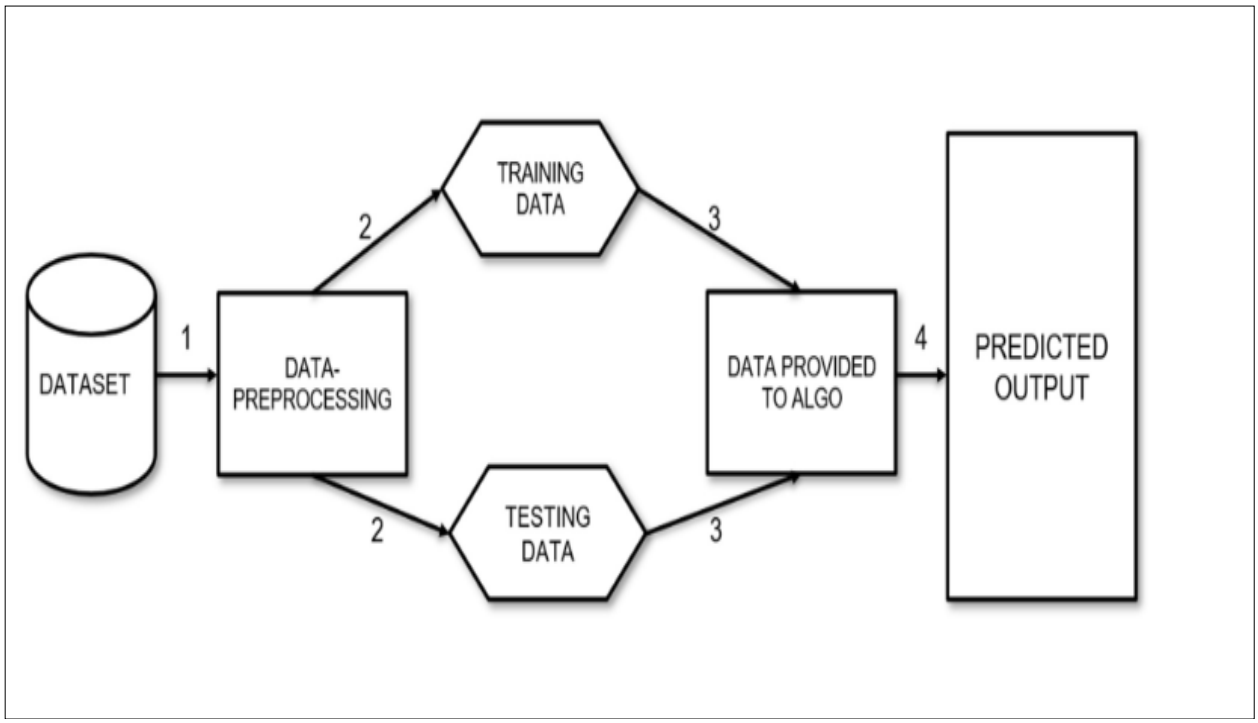

Fig No.3Working of Mobile Cloud Computing

A. Algorithm

1. Open the local host website
2. Create an account if new user or login in if already have account.
3. Select the type of attack you want to apply.
4. Capture network packets using Wireshark.
5. After the attack is done stop capturing packets.
6. Filtered the captured data packets
7. Send the filtered packets to trained ANN algorithm.
8. Then click on the detect type of attack button.
9. The algorithm will analyze the packets and give type of attack happened on the system as output.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205078 |

***B. Use Case** Diagram:*

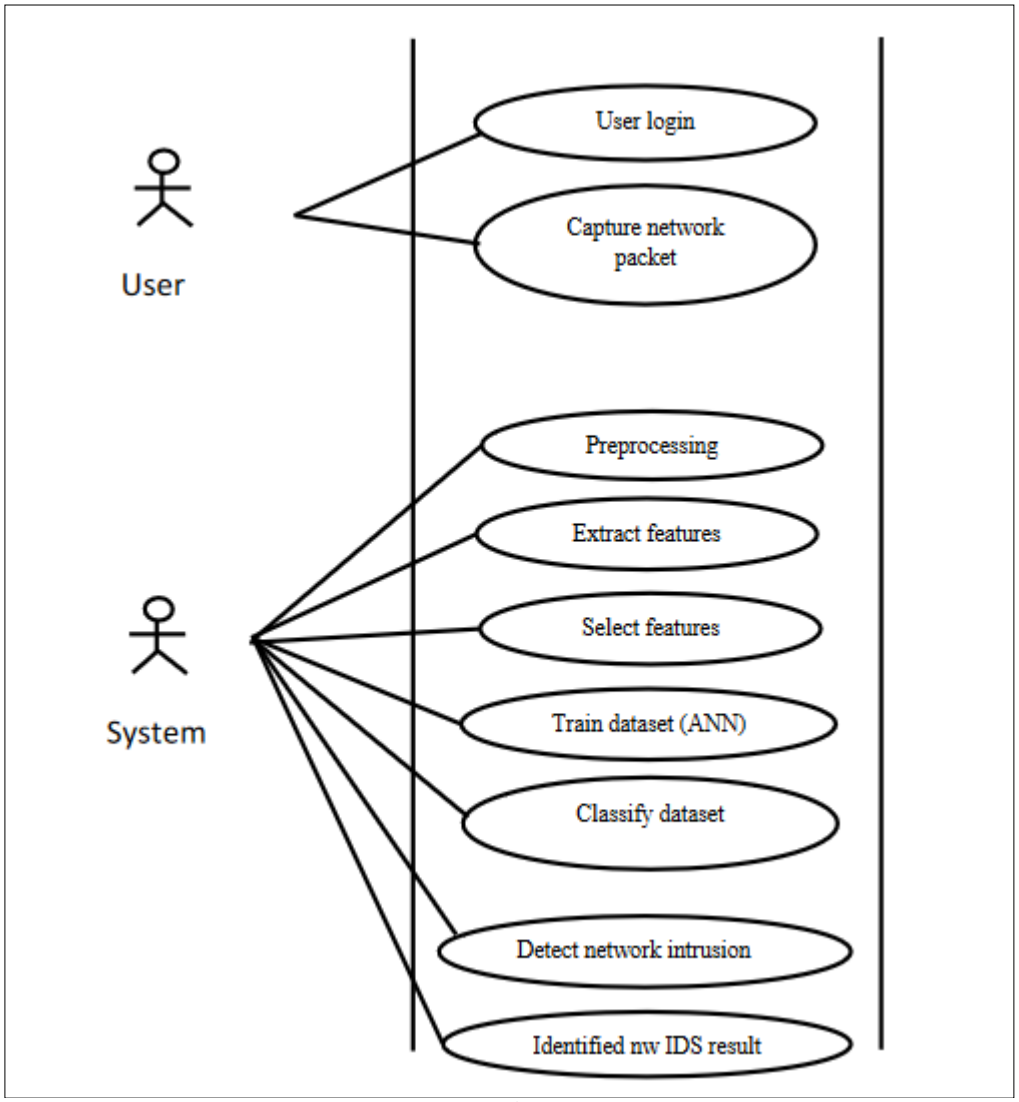

Fig No.4 User Case Diagram of Mobile Cloud Computing

C. Details Of Hardware And Software

Software Requirements

- Windows OS
- Anaconda
- Python
- Wireshark
- MySQL
- Spyder

Hardware Requirements

- Processor: Intel i3 or above
- 4 GB RAM or more
- 4 GB Graphics Card

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205078 |

Software Used

1. Anaconda:

It is a free and open-source distribution of the Python and R programming languages for data science and machine learning related applications (large-scale data processing, predictive analytics, scientific computing), that aims to simplify package management and deployment. Package versions are managed by the package management system conda. The Anaconda distribution is used by over 6 million users, and it includes more than 250 popular data science packages suitable for Windows, Linux, and MacOS.

2. Python:

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built-in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue languageto connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, whichencourages program modularity and code reuse. The Python interpreter and the extensive standardlibrary are available in source or binary form without charge for all major platforms, and can be freelydistributed.

3. Wireshark:

Wireshark lets the user put network interface controllers into promiscuous mode (if supported bythe network interface controller), so they can see all the traffic visible on that interface includingunicast traffic not sent to that network interface controller's MAC address. However, when capturingwith a packet analyzer in promiscuous mode on a port on a network switch, not all traffic through theswitch is necessarily sent to the port where the capture is done, so capturing in promiscuous mode isnot necessarily sufficient to see all network traffic. Port mirroring or various network taps extendcapture to any point on the network. Simple passive taps are extremely resistant totampering.

4. MySQL:

MySQL is well known as the world's most widely used open-source database (back-end). It is themost supportive database for PHP as PHP-MySQL is the most frequently used open-source scriptingdatabase pair. The user-interface which WAMP, LAMP and XAMPP servers provide for MySQL iseasiest and reduces our work to a large extent.

5. Spyder:

Spyder is an open-source cross-platform integrated development environment (IDE) for scientificprogramming in the Python language. Spyder integrates with a number of prominent packages in thescientific Python stack.Spyder is extensible with first-party and third-party plugins,[7] includes support for interactive toolsfor data inspection and embeds Python-specific code quality assurance and introspection instruments,such as Pyflakes, Pylint[8] and Rope. It is available cross-platform through Anaconda, on Window

Advantages

1. It improves the accuracy of result.
2. It analyzes the result of to identify most network traffic, reduces the workload.
3. Machine learning method is shows that it identifying network traffic by feature selection and extraction using ANN method and classify the dataset using SVM.
4. The benefit of the usage of machine learning is that it's going to confirm whether or not a code or a document is malicious or now no longer throughout a} very little bit of time even as not the requirement of analytic it in the course of a sandbox to carry out the analysis.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205078 |

V. CONCLUSION AND FUTURE SCOPE

The intrusion detection system we have created using machine learning algorithm ANN will provide a great security system to the computer systems of an organization or local personal computer against intrusion attacks. The use of machine learning algorithm can outperform other intrusion detection system without machine learning algorithms and can provide results for detecting intrusion with great efficiency and accuracy.

The intrusion detection system we have created using machine learning algorithm ANN will provide a great security system to the computer systems of an organization or local personal computer against intrusion attacks. The use of machine learning algorithm can outperform other intrusion detection system without machine learning algorithms and can provide results for detecting intrusionwith great efficiency and accuracy.

REFERENCES

1. Syam Akhil Repalle1, Venkata Ratnam Kolluru, 21 Student, Department of Electronics and Communication Engineering, Koneru Lakshmaiah Educational Foundation, Andhra Pradesh, January 2016
2. Md Nasimuzzaman Chowdhury and Ken Ferens, Mike Ferens1Department of Electrical and Computer Engineering University of Manitoba Winnipeg, Manitoba,Canada February 2020
3. Riyaz ahmed A. Jamadar*Department of Information Technology, AISSMS Institute of Information Technology /Savitribai Phule Pune University, march 2019
4. Jiyeon Kim , Jiwon Kim , Hyunjung Kim Minsun Shim and Eunjung Choi. 1 June 2020.
5. Bida Seraphim, Shreya Palit, Kaustubh Srivastava, E Poovammal, Implementation of Machine Learning Techniques applied to the Network Intrusion Detection System, June 2019.
6. Peng K, Leung VC, Huang Q. Clustering approach based on mini batch Kmeans for intrusion detection system over Big Data. IEEE Access. 2018.
7. Peng K. et al. Intrusion detection system based on decision tree over Big Data in fog environment. Wireless Commun Mob Comput. 2018. https://doi.org/10.1155/2018/4680867.
8. Belouch M, El Hadaj S, Idhammad M. Performance evaluation of intrusion detection based on machine learning using Apache Spark. Procedia Comput Sci. 2018;127:1–6.
9. Manzoor MA, Morgan Y. Real-time support vector machine based network intrusion detection system using Apache Storm. In: IEEE 7th annual information technology, electronics and mobile communication conference (IEMCON), 2016. Piscataway: IEEE. 2016; p. 1–5.
10. Vimalkumar K, Radhika N. A big data framework for intrusion detection in smart grids using Apache Spark. In: International conference on advances in computing, communications and informatics (ICACCI), 2017. Piscataway: IEEE; 2017. p. 198–204.
11. L. Yadav and A. Ambhaikar, "Feasibility and Deployment Challenges of Data Analysis in Tele-Healthcare System," 2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI), Raipur, India, 2023, pp. 1-5, doi: 10.1109/ICAIIHI57871.2023.10489389.
12. L. Yadav and A. Ambhaikar, "Approach Towards Development of Portable Multi-Model Tele-Healthcare System," 2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI), Raipur, India, 2023, pp. 1-6, doi: 10.1109/ICAIIHI57871.2023.10489468.
13. Lowlesh Yadav and Asha Ambhaikar, Exploring Portable Multi-Modal Telehealth Solutions: A Development Approach. International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC), vol. 11, no. 10, pp. 873–879, Mar. 2024.11(10), 873–879, DOI: 10.13140/RG.2.2.15400.99846.
14. Lowlesh Yadav, Predictive Acknowledgement using TRE System to reduce cost and Bandwidth, March 2019. International Journal of Research in Electronics and Computer Engineering (IJRECE), VOL. 7 ISSUE 1 (JANUARY- MARCH 2019) ISSN: 2393-9028 (PRINT) | ISSN: 2348-2281 (ONLINE).
15. Lowlesh Yadav and Asha Ambhaikar, “IOHT based Tele-Healthcare Support System for Feasibility and performance analysis,” Journal of Electrical Systems, vol. 20, no. 3s, pp. 844–850, Apr. 2024, doi: 10.52783/jes.1382.

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

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