Network Intrusion System Using The K Nearest Neighbor Network Algorithm

Bada, Oluwafunke A. & OBARAYI, Z.O.
The Federal Polytechnic, Ilaro,
Computer Science Department, federal polytechnic, ilaro, Nigeria.
[1]oluwafunke.bada@federalpolyilaro.edu.ng

Abstract
Technological development has helped improve the lives of humanity. The proliferation of technology is significantly helping to transform manual operations and improve their performances. As indicated with any of some other structure, there are potentials for vulnerability and intrusion of such spaces. Hence, there is a need to create a more technological way to detect and possibly find solutions to these breaches of technology. The machine learnings have had its own applications in both image and speech recognition, general prediction and even online fraud detection. The features and the structure of machine learning concepts is utilized to solve issues of problems of network intrusion and vulnerabilities. This intrusion detection system strives to address widespread inefficiencies associated with the traditional manual method of intrusion detection. This system aims to conceptualize and create a dedicated intrusion detection system by solving the problem of the Denial of Service, attack that consumes Limited and non-renewable resources. This system aims to design a network intrusion system using the K nearest neighbor network algorithm that achieves high detection of accuracy and with a stand zero attacks.

Keyword: Denial of service, Machine Learning Intrusion, k nearest neighbor, vulnerability.

1.0 Introduction

Technological development has helped improve the lives of humanity. The proliferation of technology is significantly helping to transform manual operations and improve their performances. Nowadays, we live in a data based dynamic world which is connected through set of networks. As indicated with any of some other structure, there are potentials for vulnerability and intrusion of such spaces. Hence, there is a need to create a more technological way to detect and possibly find solutions to these breaches of technology (Mohajan, 2021).

The machine learnings has had its own applications in both image and speech recognition, general prediction and even online fraud detection. We are going to utilize the features and the structure of machine learning concepts to solve issues of problems of network intrusion and vulnerabilities (Mazhar et. al., 2023).

K-Nearest Neighbor is a classifier used in data mining. Fix and Hodges proposed the supervised KNN classifier in 1951.By determining the k nearest neighbors and computing the Euclidean Distance, the output of the target variable is anticipated. According to Jabbar et al. (2013), it is a non-parametric classification method that makes no assumptions about the underlying data.

One kind of attack or intervention that takes place inside a system is called intrusion. IDS is a piece of software or an application designed to monitor and analyze system network traffic and keep hackers out. IDS is a hardware and software combination that performs intrusion detection. IDS can determine whether a computer network or system has a violation or an attack by gathering and examining a few important aspects of the system.

2.0 Literature Review

According to Mohajan (2021), the development of internet and computer technology, as well as the numerous opportunities they present to further human pursuits, were the hallmarks of the third industrial revolution. Because of this, many processes that were formerly only physical are now both physical and electronic, or completely digital. Payment for goods and services, for example, may now be made with little to no physical presence of the buyer and seller thanks to the availability of online vendors, which enables both parties to conduct business online. Additionally, thanks to digital technology, banking services are no longer always provided in a physical hall; instead, they can be accessible by merely pressing buttons or swiping across the interfaces of computer devices that are connected to other devices.

A computer network is defined as a network of computing systems that can share assets in along with exchanging information and data (Amazon, 2023). But while computer networks and the internet in general have greatly benefited humanity, the criminal potential of these technological developments keeps growing. Every day, criminal elements saturate the internet and

other intra-networks in pursuit of victims. For a variety of motivations, including retaliation, espionage, ransom, competition, ego boosts, pure enjoyment, etc., these malevolent individuals frequently prey on victims. (Bandakkanavar, 2023); (Otienofedi, 2023).

In order to undermine the efficient operation of networks of computers for illicit purposes, various threats and attacks have been used (Orukpe et al., 2013). These network threats and attacks, which include distributed denial of service (distributed denial-of), denial of service (DoS), worms, Trojan horses, viruses, injection attacks, and more (Shruti, 2023), aim to undermine the availability, security, and reliability of the internet and computer networks (Osa, 2022). Given the ongoing increase of internet users and the development of the internet of things (IoT), it is more important than ever to offer treatments that can identify dangers and attacks that could jeopardize information networks (Mazhar et al., 2023).

In the field of cyber security, intrusion detection systems (IDS) are a common intervention. Simply described, an intrusion detection system is a combination of software and hardware used to identify intrusions or attacks in networks of computers (Devi & Abualkibash, 2019). In order to find and report intrusions based on preconfigured detection flags, an IDS, or intrusion detection system, analyzes network traffic (Chudasma, 2020). A single computer system or an enormous network infrastructure could be the typical IDS (Sangfor Technologies, 2023). The two categories of intrusion detection systems (IDS) that are based on the degree of occupancy are host-based IDS (HIDS) and network intrusion detection systems (NIDS). An impressive example of HIDS is a computer system that monitors important operating system files and protects against both internal and external compromise (Sangfor Technologies, 2023).

Additionally, the way that IDSs detect also leads to the creation of new IDS classifications, such as anomaly-based IDS and signature-based IDS (SIDS). Typically, a signature-based intrusion detection system (IDS) looks for particular patterns in data traffic, such as byte or instruction sequences, to identify intrusions. After that, it compares them to a database of malicious code signatures that have already been identified (Velimirovic, 2023). This detecting method is comparable to how antivirus software behaves. Although SIDSs are highly efficient against known attack signatures, their capacity to identify attacks without a pre-existing signature is constrained (Velimirovic, 2023). Within the field of information security, such attacks are referred to as zero-day attacks.

Because malware was developing so quickly, anomaly-based intrusion detection systems, or AIDSs, were introduced. As opposed to SIDS, which use a database of known attack signatures, such IDSs can monitor network data traffic by comparing it with a predetermined baseline that is regarded as normal behavior. An AIDS looks for unusual activity in all areas of the network layer, including devices, ports, bandwidth, protocols, and more. It is possible to create extremely intelligent anomaly-based intrusion detection systems (AIDSs) by using machine learning in their construction.

3.0 Methodology

The network intrusion detections system developed in this research study is based on existing implementation infrastructures and also modeled according to the existing mode of intrusion detection systems using the k nearest neighbor. This system was designed and developed to serve as a cyber-security technology that overseer network traffic for high potential security threats and alerts administrators to potential incidents. The system is designed to alert and detect on unauthorized access, misuse, or other malicious activities that may compromise the security of a network.

This network intrusion detection system implements the k-nearest neighbors (KNN) algorithm. The k-nearest neighbors (KNN) algorithm is a non-parametric, supervised learning classifier, which uses proximity to make classifications or predictions about the grouping of an individual data point. In this study, network data was collected and pre-processed, a machine learning intrusion detection system prototype was built, and analysis was conducted. Data will be collected from Kaggle, the data will be pre-processed and normalized. Finally, using the normalized train and test sets, the IDS prototype will be tested and an analysis of the results will be performed. The software development life cycle (SDLC) for this system is the waterfall model. The waterfall approach is called a linear or sequential method because it involves different phases, each leading to the next. The sequential approach begins at the system level and progresses through analysis, design, coding, testing, and support. (Choudhury, 2015). This software development framework was considered for the successful process of structuring, planning,

implementing and controlling the developed combination of hardware and software-based network intrusion detection system. This approach helps to identify potential risks and vulnerabilities in the network, and also provides measures to prevent them. In addition, it ensures the compliance of the system with the relevant regulations and ensures its security. This is to ensure the quality of software and hardware design within the requirements of the software development life cycle. The development and coding architecture of this network intrusion detection system includes both hardware and software system design. A HTML, CSS, and JavaScript interface is designed and implemented.

Flowchart of the system

The flowchart that describes the system is described in different sections depending on the interface and the operations that each user of the system can perform. The application was developed in several stages, as shown in the figure below. Milestones include all design and development efforts that have been completed. The following sections describe each step in detail

.

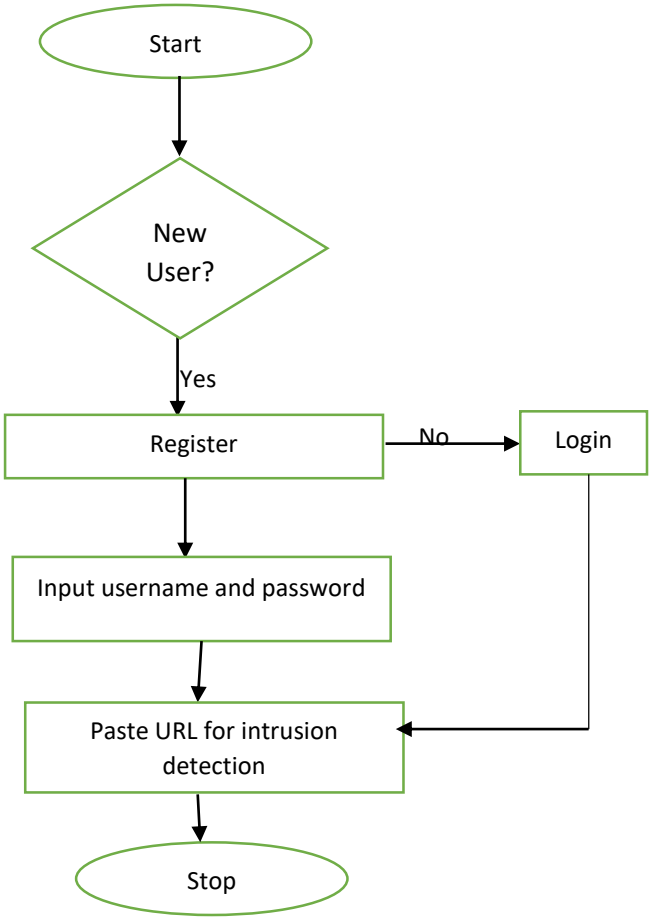

Figure 1.0 Flowchart Diagram

The use case diagram in Figure 1.1 describes the actions that can be performed by the system user.

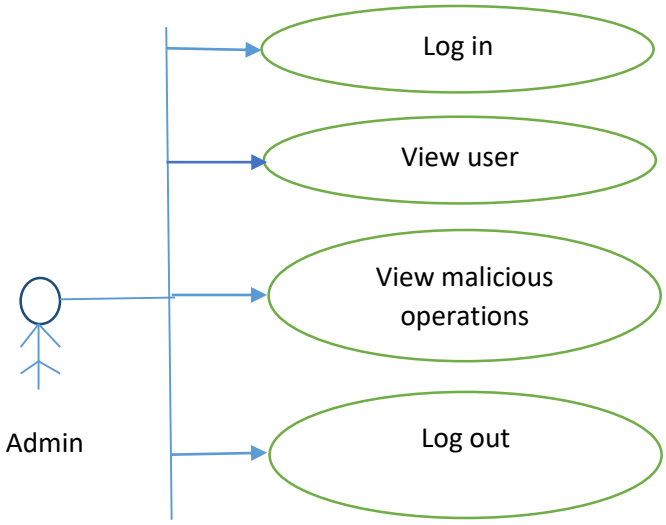

Figure 1.1 Admin Use Case Diagram

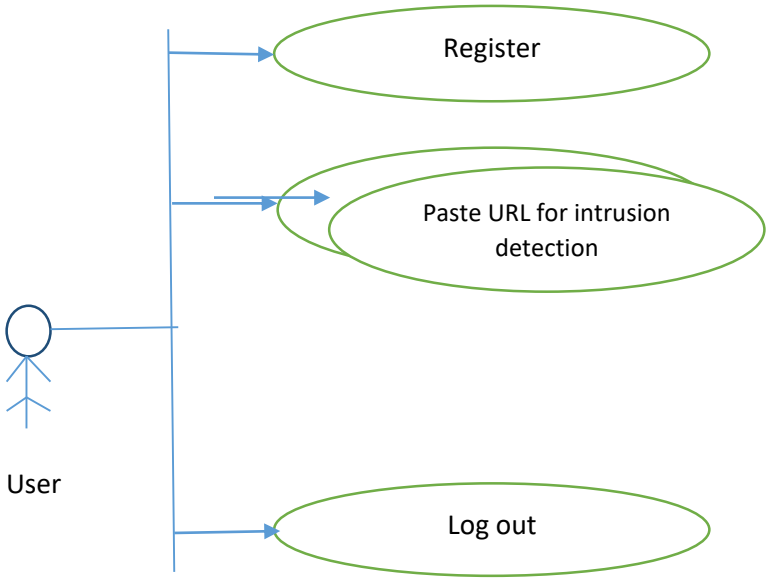

Figure 1.2 User's Use Case Diagram

4.0 Result and Discussion

4.1 System Implementation

The implementation of a Network Intrusion Detection System (NIDS) aims to monitor and analyze network traffic for signs of suspicious activity or potential security breaches. This system is designed to identify unauthorized access, malware, and various types of attacks such as Denial of Service (DOS), phishing, and brute force attacks.

The Graphical User Interface (GUI) of the Network Intrusion Detection System (NIDS) is designed to provide network administrators with an intuitive and user-friendly platform to monitor, analyze, and respond to network threats. The interface ensures that critical security information is easily accessible, enabling real-time tracking and management of potential intrusions. Below is a description of the main components that typically be feature in the GUI of the developed system:

The Landing Page: This page gives an overview of the system. It allows the users of the system to easily navigate to all pages of the system.

Figure 1.3. The landing page

The Registration Page: The registration page allows new users to create an account on the website. They provide information such as name, email address and password to register on the website.

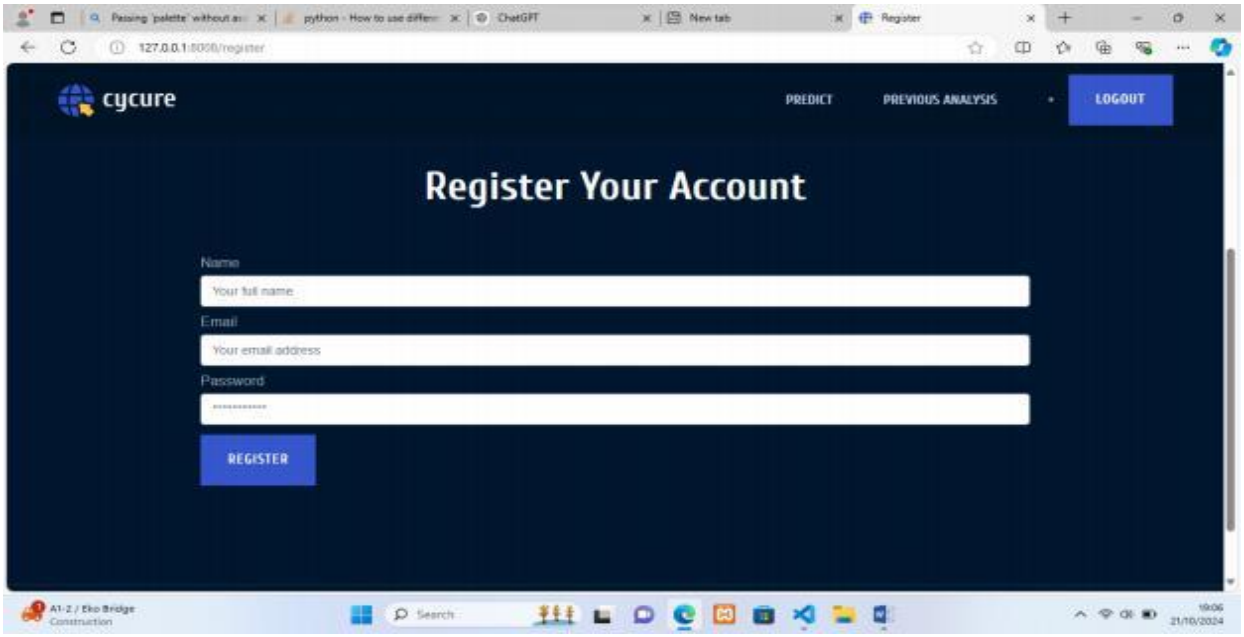

Figure 1.4 the registration page

***The Log In Page*:** The login page allows existing users to access their account on the website. They provide information such as email address and password in order to access their accounts whenever they want to login to the system.

Figure 1.5. The login page

The Intrusion Detection Page: This page allows users to detect intrusion by providing the websites' frequency, range, bytes, payload size, interval, and repeated bytes count for accurate prediction result.

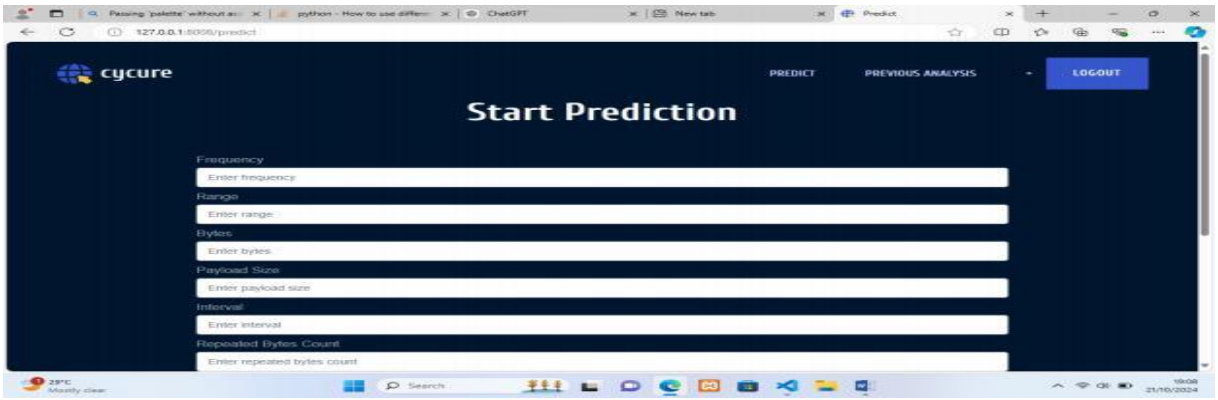

Figure 1.6.. The intrusion detection page

The Intrusion Detection Result Page: This page displays the result of the intrusion detection prediction.

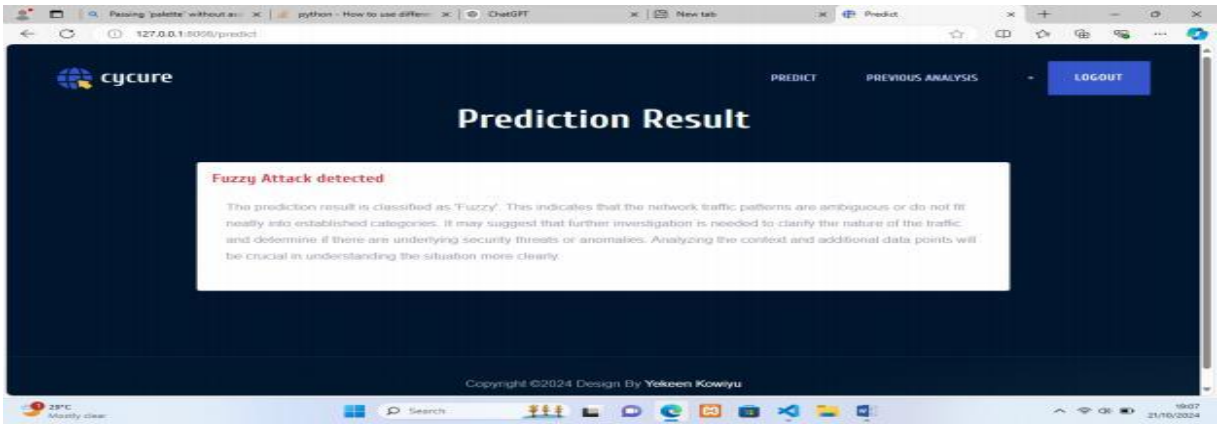

Figure 1.7. The intrusion detection result page

***The Ids History Page*:** This page displays the lists of the previously done predictions, their details and also the result of each of them.

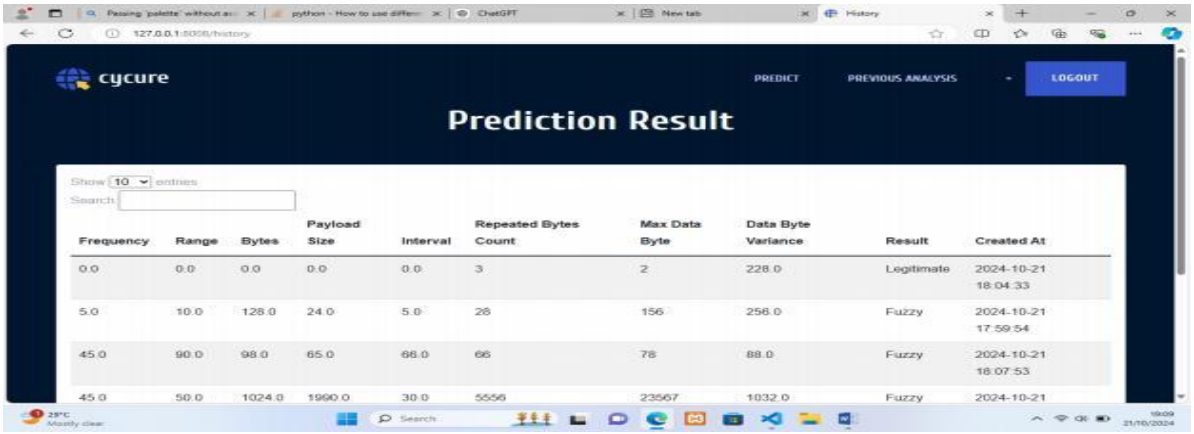

Figure 1.8. The IDS history page

5. Conclusion and Future Works

This Network Intrusion Detection System (NIDS) project focuses on the development of a system designed to monitor and analyze network traffic for suspicious activities that could signal security breaches or cyber-attacks. Using a combination of signature-based detection (which matches known patterns of malicious behavior) and anomaly-based detection (which identifies deviations from normal traffic behavior), the system aims to detect various forms of network intrusions, such as unauthorized access, malware attacks, and data breaches. The KNN Machine learning model and rule-based technique was used to identify threats in real-time, enabling network administrators to respond promptly to any potential security incidents.

The project involved data collection from network traffic logs, preprocessing of the data, feature extraction, and training KNN machine learning algorithm to classify normal and suspicious traffic. The system was evaluated using performance metrics such as accuracy, detection rate, false positives, and false negatives. It also included a user interface for real-time monitoring and alert generation.

The Network Intrusion Detection System developed in this project proves to be an effective solution for identifying suspicious network activity in real-time. By combining machine learning techniques with signature-based methods, the system can detect a wide variety of known and unknown threats, offering a more robust security solution.

The real-time monitoring capability ensures that network administrators are alerted promptly, enabling faster incident response and mitigation. However, as cyber-attacks evolve, the system will require continuous updates and training on new attack patterns to maintain its efficacy.

References

Aakanksha, C. (2021). Security Issues of Firewall. *International Journal of P2P Network Trends and Technology (IJPTT).* Vol. 6, Issue 1, January to February.

Aas, K., and Eikvil, L. (2020). An Evaluation of Statistical Approaches to Text Categorization, Technical Report CMU-CS-97-127, Computer Science Department, Carnegie Mellon University.

Abuh, T.O., and Orukpe, P.E. (2020*). Computer security and privacy in wireless local area network in Nigeria, Int. J. Eng. Res.* Afr. Pg. 23–33. Trans Tech Publications, Switzerland.

Amazon (2023). 'What is computer networking?' available at: www.amazon.com.

Amreen, S., and Jabbar, M. A. (2019)." intelligent network intrusion detection system using data mining techniques" IEEE explore 2019.

Bandakkanavar, R. (2023). 'Causes of cybercrime and preventive measures.

Choudhury, K., (2015). *The Art of Agile Practice: A Composite Approach for Projects and Organizations*. CRC Press. pp. 56–59. ISBN 9781439851197.

Chudasma, P. (2020). 'Network intrusion detection system using classification techniques in machine learning' [MSc. Data Analytics] at Dublin Business School (2020).

Devi, R.R., and Abualkibash, M. (2019). Intrusion detection system classification using different machine learning algorithms on KDD-99 and NSL-KDD datasets - *a review paper, Int. J. Comput. Sci. Inf. Technol. 11 (3) (2019) (IJCSIT) VolJune.*

G´eron, A. (2019). 'Hands-on machine learning with scikit-learn, Keras, and TensorFlow' O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472 (2019).

Haripriya, L., and Jabbar, M. A. (2020)." A Novel intrusion detection system using ANN and feature subset selection", *international journal of engineering and technology.*