AI SECURITY FRAMEWORK

[1]Mrs.V Soujenya,[2]Uppuluri Sai Karthik,[3]Tiyyagura Praneeth Reddy[4]Gayala Nithin

[1]Assistant Professor, Dept. of CSE-Cyber Security, Geethanjali College of Engineering and Technology Cheeryal (V), Keesara (M), Medchal(D), Hyderabad, Telangana 501301,

soujenyav.cse@gcet.edu.in

[2, 3, 4, BTech] Student, Dept. of CSE-Cyber Security, Geethanjali College of Engineering and Technology Cheeryal (V), Keesara (M), Medchal(D), Hyderabad, Telangana 501301,

20r11a6248@gcet.edu.in,20r11a6256@gcet.edu.in,21r15a6204@gcet.edu.in

ABSTRACT:

The creation of effective cyber security measures is required due to the swift rise and spread of cyber threats. Sophisticated machine learning (ML) and artificial intelligence (AI) algorithms are helping to coordinate more and more modern cyber-attacks by using data to mimic and learn from user behavior. In order to counter these types of threats, it becomes necessary to implement systems that are capable of independently and expertly identifying attack patterns. Moreover, the development of cyber defenses has tremendous opportunities thanks to ML and AI technology. Within this framework, our research activities span three primary areas:-

(1) Intrusion detection in cyber security using ML.

(2)Detection of Phishing URL's using ML.

(3) Detection of XSS scripts using ML.

(4)Detection of SQL queries using ML.

Keywords: Cyber security, sophisticated machine learning (ML).

I INTRODUCTION

In the face of an ever-expanding landscape of cyber threats fueled by sophisticated machine learning (ML) and artificial intelligence (AI) algorithms, the imperative to fortify cyber security measures has reached unprecedented heights. This project represents a proactive response to the swift evolution and proliferation of cyber threats, strategically employing cutting-edge technologies to not only counteract the growing complexity of modern cyber-attacks but also to harness the potential of AI and ML for the development of robust cyber defenses. The research is structured around three pivotal areas, each addressing critical aspects of cyber security in the contemporary digital environment. These areas include the application of ML for intrusion detection, the classification and generation of cyber security-relevant strings such as passwords and phishing URLs. Through a comprehensive exploration of these domains, the project aims to contribute innovative solutions to the challenges posed by the dynamic and sophisticated nature of cyber threats in the 21st century.

The objectives of this project are:

- Intrusion detection in cyber security using ML tools.
- Classification and generation of strings that are relevant for cyber security, in phishing URLs.
- Detecting XSS and SQL Injection attacks

II. LITERATURE SURVEY

1. Buczak and Guven [1] in this paper, they focus on machine-learning methods and their applications to intrusion detection. Algorithms such as Neural Networks, Support Vector Machine and Bayesian Networks are described in detail. However, major ML methods such as clustering, and Gradient Boosting are not included. Their paper focuses on network intrusion detection. Through a wired network, attackers must pass multiple layers of firewall and operating system defenses or gain physical access to the network. However, any node can be a target in a wireless network; thus, the network is more vulnerable to malicious attacks and is more difficult to defend than are wired networks. Limitation of this paper is, they have use only 3 machine learning methods and didn’t use major machine learning methods.

2. Samson Ho and Rabindra [2]The major objective of this paper is to propose a Convolutional Neural Network-based Intrusion Detection System (IDS) for enhancing internet security. The proposed

IDS paradigm categorizes all network packet traffic into good and malicious kinds in order to detect network intrusions. The Canadian Institute for Cybersecurity's CICIDS2017 dataset was used to train and validate the suggested model. All aspects of the model have been evaluated, including over all accuracy, attack detection rate, false alarm rate, and training overhead. Three other well-known classifiers have been compared with the recommended model to see how effective they are. Limitation of this paper is, the effectiveness of the proposed IDS might be limited to known attacks present in the training dataset. The model may struggle to generalize well to new, unseen types of attacks that emerge after the training phase.

3. Mohamed M and Abhishek [3]this article suggests using IDS with two layers. The first layer categorizes the network connection based on the service being used. Following that, a minimal set of features that improve the detection precision of malicious activity on that service are found. The second layer uses those features to categorize each network connection as an attack or regular activity using the pattern recognition technique. The normal behavior model and the attack behavior model are two multivariate normal statistical models that are produced during the training phase. The two multivariate normal statistical models are employed in the testing and operating phases to classify a network connection into attack or normal activity using a maximum likelihood estimate function. Limitations of this paper is, while pattern recognition techniques are powerful, they may struggle with highly complex or evolving attack patterns. If attackers employ sophisticated evasion techniques or constantly change their strategies, the pattern recognition models may not be able to keep up.

4. M. Hall, E. Frank, J. Holmes [4] In this paper they put forth an extensive survey about all the major e-mail filtering and ML techniques that can be used to classify and recognize phishing emails from normal ones. Their paper provides a thorough analysis of the most recent studies on phishing attempts and a comprehensive summary of the different approaches used to identify and categorize these malicious emails. The assessment covers a broad spectrum of methods, such as sophisticated machine learning algorithms, heuristic approaches, and conventional rule-based filtering. Moreover, they perform a comparative analysis, assessing the efficiency of every method to clarify their individual advantages and disadvantages. This comparative analysis offers insightful information to our project. Limitations of this paper is, Phishing techniques are constantly evolving, and the dataset used for analysis may not capture the most recent tactics employed by attackers. The effectiveness of the methods might vary over time due to changes in phishing strategies.

III SYSTEM ANALYSIS

EXISTING SYSTEM

The absence of a well-defined existing system for the proposed project necessitates a meticulous examination of the current literature to delineate the goals and objectives of the research endeavor. In the absence of a suitable structure tailored to the unique challenges posed by the evolving landscape of cyber threats, a comprehensive review will serve as the foundational step towards establishing a coherent framework for the project. Recognizing the imperative of an informed and strategic approach, the research endeavors to contribute to the advancement of cyber security by proposing innovative solutions grounded in state-of-the-art research. The project aims to identify gaps, challenges, and opportunities in the field, laying the groundwork for the development of a robust and effective cyber security framework tailored to address the contemporary complexities of cyber threats facilitated by machine learning and artificial intelligence algorithms

Limitations of Existing System

- Accuracy less
- Low efficiency

PROPOSED SYSTEM

The objective of the proposed system is to revolutionize cyber security defenses in response to the escalating sophistication of cyber-attacks leveraging machine learning (ML) and artificial intelligence (AI) techniques. The system is strategically designed to address this imperative through focused research efforts in three key domains. Firstly, the system aims to enhance cyber security efficacy by leveraging ML techniques for intrusion detection, thereby bolstering the ability to identify and thwart evolving cyber threats. Secondly, the system incorporates sophisticated categorization and generation algorithms to fortify cyber security measures pertaining to critical elements such as phishing URLs, XSS scripts and SQL queries aiming to rectify vulnerabilities associated with these integral components. Effectively mitigating the challenges posed by increasingly complex cyber-attacks in the era of ML and AI technologies.

Proposed system Advantages:

● Comprehensive Security Coverage:- By incorporating modules to detect SQL injection, phishing URLs, cross-site scripting (XSS), and malicious network packets, the website provides comprehensive coverage against various common security threats encountered in web applications and networks.

• Automation:- By leveraging machine learning models, the website automates the process of security threat detection. This reduces the need for manual intervention and accelerates the analysis process, enabling users to make faster decisions to mitigate potential risks.

• Continuous Improvement:- As new machine learning algorithms and techniques are developed, the website can be updated to incorporate the latest advancements in security threat detection. This ensures that the website remains effective and relevant in an ever-evolving threat landscape

IV IMPLEMENTATION

Architecture:

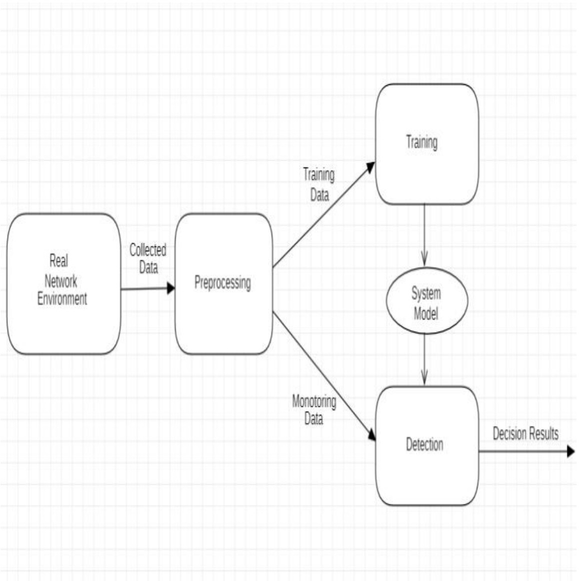


Fig-1. Architectures of the system model

The diagram depicts the architecture of our system. The first stage is data collection, where the data can be gathered from the real network or environment. This data is then preprocessed where we select the important features from the data to get it ready for training. Then, the data is used to train the system model. The final step is decision where the model gives the prediction to the new data provided by the user.

MODULES

Malicious SQL Query Detection Module:

This module utilizes machine learning algorithms to identify and block malicious

SQL queries that attempt to exploit vulnerabilities in the database. It analyzes incoming SQL queries in realtime, flagging those that exhibit patterns indicative of malicious intent, such as SQL injection attacks.

XSS Script Detection Module:

This module employs machine learning techniques to detect cross-site scripting (XSS) attacks, which involve injecting malicious scripts into web pages viewed by other users. It analyzes the content of incoming requests and responses, identifying and neutralizing XSS payloads to prevent client-side attacks.

Phishing URL Detection Module:

Description: Leveraging machine learning models, this module identifies and blocks URLs associated with phishing attacks, which aim to deceive users into providing sensitive information. It evaluates URLs in real-time, assessing various features and indicators to determine their likelihood of being malicious.

Malicious Packet Data Detection Module:

Description: This module employs machine learning algorithms to analyze network traffic and identify patterns indicative of malicious activity within packets. It inspects packet headers and payloads, looking for anomalies, signatures of known threats, or behaviors consistent with cyber-attacks.

Process:

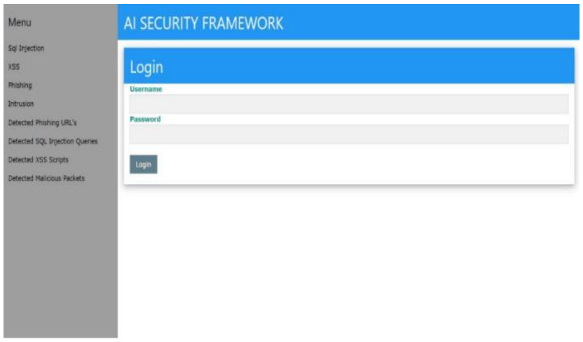
The "AI Security Framework" project encompasses a multifaceted approach to safeguarding web applications against various cyber threats. Leveraging machine learning (ML) techniques, the framework integrates four distinct modules tailored to detect different types of malicious activities. Firstly, the SQL injection module utilizes a Logistic Regression model to identify potentially harmful SQL queries inputted by users, mitigating the risk of database manipulation attacks. Secondly, the Cross-Site Scripting (XSS) module employs an Artificial Neural Network (ANN) model to detect and prevent XSS attacks, ensuring the integrity of web content by identifying and neutralizing malicious scripts. Thirdly, the phishing detection module utilizes another ANN model to scrutinize user-inputted URLs, identifying phishing attempts to protect users from falling victim to fraudulent websites. Lastly, the intrusion detection module employs three distinct models - K-Nearest Neighbors (KNN), Bernoulli Naive Bayes (BNB), and Decision

Tree - to analyze packet data for signs of malicious activity, bolstering network security by swiftly identifying potential threats. Each module within the AI Security Framework incorporates carefully curated datasets sourced from platforms like Kaggle, providing a diverse range of training samples to enhance model accuracy and robustness. By utilizing a combination of supervised learning algorithms and domain-specific datasets, the framework achieves high precision and recall rates in detecting and mitigating cyber threats across different vectors. Moreover, the framework includes user-friendly interfaces for monitoring and reviewing detected malicious data within each module, enabling administrators to gain insights into security incidents and take proactive measures to fortify web applications against evolving threats. Overall, the AI Security Framework represents a comprehensive and proactive approach to cyber security, leveraging the power of machine learning to safeguard against a wide array of potential vulnerabilities and attacks

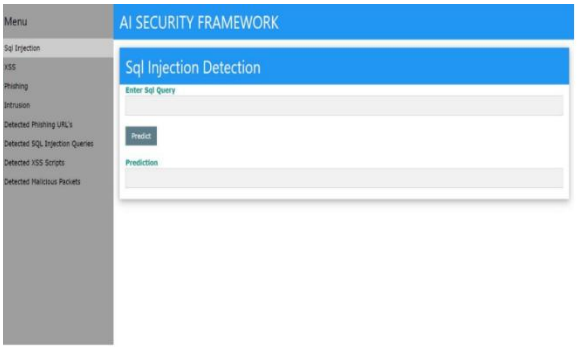
V RESULT AND DISCUSSION

Home page:

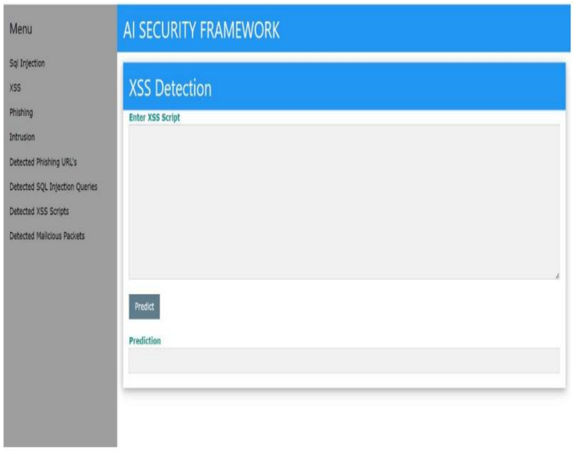
Login Page:



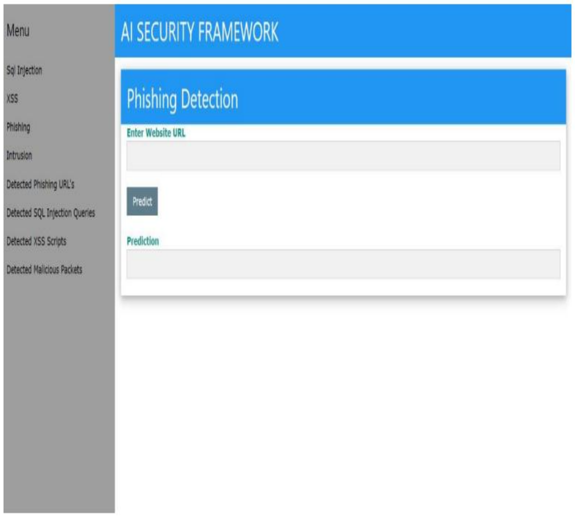
Sql injection detection:



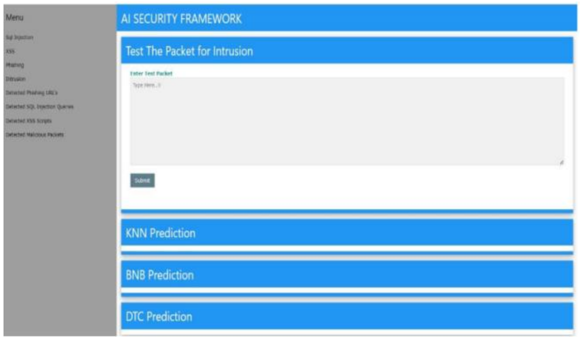
Xss Detection:



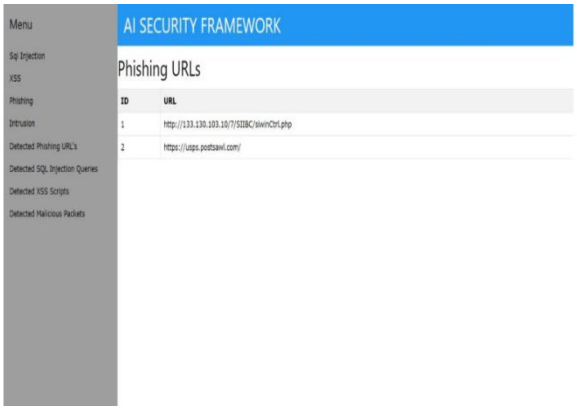
Phishing Detection:



Pocket for intrusion:



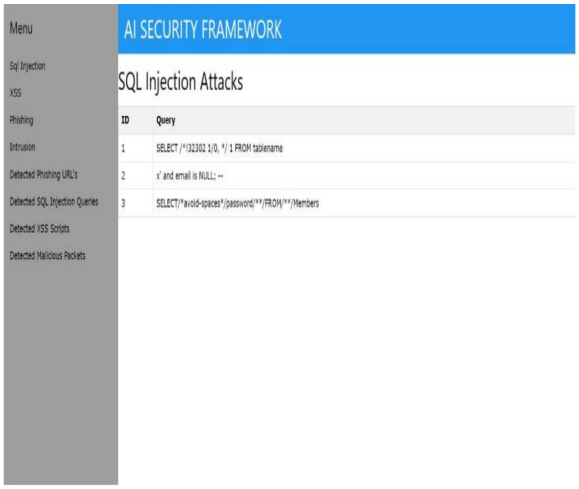
Phishing URLS:



Phishing URLs

ID	URL
1	http://133.130.103.10/7/SIIBC/siwinCtrl.php
2	https://usps.postsawl.com/

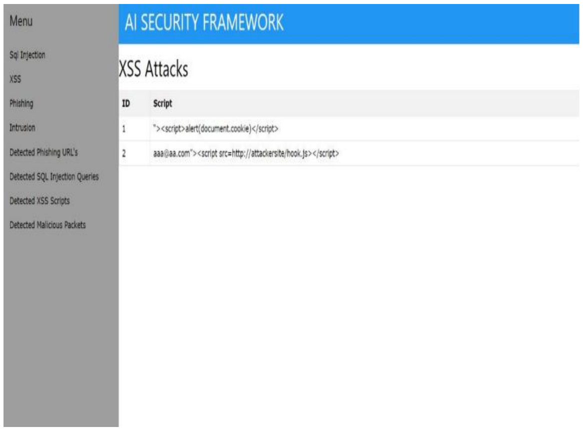
Sql injection Attacks:



SQL Injection Attacks

ID	Query
1	SELECT /*!32302 1/0, */ 1 FROM tablename
2	x' and email is NULL; --
3	SELECT/*avoid-spaces*/password/**/FROM/**/Members

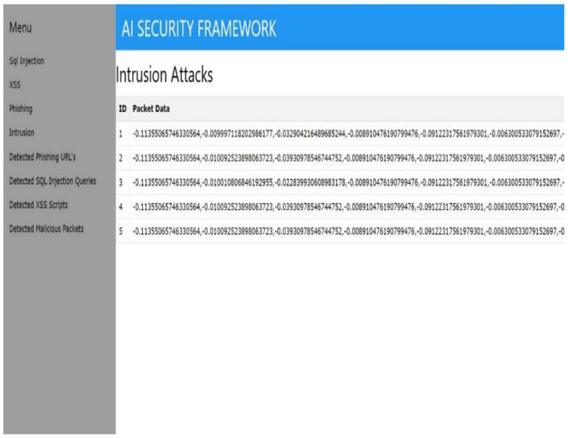
XSS Attacks:



XSS Attacks

ID	Script
1	"><script>alert(document.cookie)</script>
2	aaa@aa.com"><script src=http://attackersite/hook.js></script>

Intrusion Attacks:



Intrusion Attacks

ID	Packet Data
1	-0.11355065746330564,-0.009997118202986177,-0.032904216489685244,-0.008910476190799476,-0.09122317561979301,-0.006300533079152697,-
2	-0.11355065746330564,-0.010092523898063723,-0.03930978546744752,-0.008910476190799476,-0.09122317561979301,-0.006300533079152697,-0
3	-0.11355065746330564,-0.010010806846192955,-0.022839930608983178,-0.008910476190799476,-0.09122317561979301,-0.006300533079152697,-
4	-0.11355065746330564,-0.010092523898063723,-0.03930978546744752,-0.008910476190799476,-0.09122317561979301,-0.006300533079152697,-0
5	-0.11355065746330564,-0.010092523898063723,-0.03930978546744752,-0.008910476190799476,-0.09122317561979301,-0.006300533079152697,-0

VI CONCLUSION

The AI Security Framework represents a significant step forward in safeguarding web applications against a myriad of cyber threats. Through the integration of specialized modules employing machine learning algorithms, the framework demonstrates its effectiveness in detecting and mitigating various attack vectors such as SQL injection, XSS scripting, phishing attempts, and network intrusions. By employing distinct models tailored to each threat type, the framework ensures comprehensive coverage and robust protection for web-based systems, enhancing overall cyber security posture. One of the standout features of this website is its user-friendly interface, which makes it accessible to a wide range of users, from individuals concerned about personal cyber security to large organizations seeking comprehensive network protection. The automation capabilities further streamline the security process, allowing for swift responses to emerging threats without extensive manual intervention.

Furthermore, the rigorous testing conducted during the development phase validates the reliability and accuracy of the AI Security Framework. Test cases spanning different attack scenarios and input types confirm the framework's ability to accurately classify and respond to potential security threats in real-time. The incorporation of diverse datasets sourced from reputable platforms like Kaggle ensures that the models are trained on a wide range of examples, enhancing their ability to generalize and adapt to evolving cyber threats.

Moving forward, the AI Security Framework stands poised to make significant contributions to the ongoing battle against cybercrime. As threats continue to evolve and grow in sophistication, the framework's adaptability and scalability position it as a vital tool for organizations seeking to fortify their digital assets against malicious actors. By leveraging the power of machine learning and proactive threat detection, the framework empowers administrators with the tools and insights necessary to stay one step ahead of cyber threats, ultimately fostering a safer and more secure online environment for all users.

FUTURE ENHANCEMENT

• Enhanced User Interface: Improve the user interface to provide more intuitive and interactive features. This could include visualizations of threat analysis results, detailed explanations of detected threats, and suggestions for remediation. Support for Custom ML Models: Allow users to train and deploy their custom machine learning models tailored to their specific security requirements. This would enable organizations to adapt the system to their unique threat landscape and data characteristics.

• Incident Response Automation: Implement automated incident response mechanisms based on the detected threats. This could include triggering alerts, blocking malicious traffic, quarantining affected resources, or initiating remediation actions.

VII REFERENCES

1. Sarker IH (2022) Smart city data science: towards data-driven smart cities with open research issues. Internet Things 19:100528

2. Sarker IH, Asif IK, Yoosef BA, Fawaz A (2022) Internet of things (IoT) security intelligence: a comprehensive overview, machine learning solutions and research directions. Mobile Netw Appl117

3. Sarker IH (2021) Machine learning: algorithms, real-world applications and research directions. SNComputSci 2(3):1–21

4. Sarker IH (2021) Cyberlearning: effectiveness analysis of machine learning security modeling to detect cyber-anomalies and multi-attacks. Internet Things 14:100393 5. Tien JM (2017) Internet of things, real-time decision making, and artificial intelligence. Ann Data Sci 4(2):149–178

6. Shi Y (2022) Advances in big data analytics: theory, algorithms and practices. Springer, Berlin

7. Sarker IH, Kayes ASM, Badsha S, Alqahtani H, Watters P, Ng A (2020) Cybersecurity data science: an overview from machine learning perspective. J Big Data 7(1):1–29

8. S´lusarczyk B (2018) Industry 4.0: are we ready? Pol J Manag Stud 17:232–248

9. Sarker IH, HasanFurhad M, Nowrozy Ra (2021) AI-driven cybersecurity: an overview, security intelligence modeling and research directions. SN ComputSci 2(3):1–18

10. Sarker IH (2022) AI-based modeling: techniques, applications and research issues towards automa-tion, intelligent and smart systems. SN ComputSci 3(2):1–20

[11] Prasadu Peddi (2015) "A review of the academic achievement of students utilising large-scale data analysis", ISSN: 2057-5688, Vol 7, Issue 1, pp: 28-35.

[12] Prasadu Peddi (2015) "A machine learning method intended to predict a student's academic achievement", ISSN: 2366-1313, Vol 1, issue 2, pp:23-37.

AUTHORS

Mrs. V Soujenya,Assistant Professor Dept. of CSE-Cyber Security, Geethanjali College of Engineering and Technology Cheeryal (V), Keesara (M), Medchal(D), Hyderabad, Telangana 501301.

Email: soujenyav.cse@gcet.edu.in

Mr. Uppuluri Sai Karthik, Dept. of CSE-Cyber Security, Geethanjali College of Engineering and Technology Cheeryal (V), Keesara (M), Medchal(D), Hyderabad, Telangana 501301.

Email: 20r11a6248@gcet.edu.in

Mr.Tiyyagura Praneeth Reddy, Dept. of CSE-Cyber Security, Geethanjali College of Engineering and Technology Cheeryal (V), Keesara (M), Medchal(D), Hyderabad, Telangana 501301.

Email: 20r11a6256@gcet.edu.in

Mr. Gayala Nithin, Dept. of CSE-Cyber Security,Geethanjali College of Engineering and Technology Cheeryal (V), Keesara (M), Medchal(D), Hyderabad, Telangana 501301.

Email: 21r15a6204@gcet.edu.in