

# DETECTING CYBER ATTACKS THROUGH MEASUREMENTS: LEARNING FROM A CYBER RANGE

<sup>1</sup>Dr.CH.V. Phani Krishna, <sup>2</sup>BADIKELA LIKITHA, <sup>3</sup>CHENNUPALLI YASHWANTH SAI,  
<sup>4</sup>EDAM BHARGAV SAI

<sup>1</sup>(Professor) ,CSE. Teegala Krishna Reddy Engineering College Hyderabad

<sup>2,3,4</sup>B,tech scholar ,CSE. Teegala Krishna Reddy Engineering College Hyderabad

## ABSTRACT

Malicious URLs pose a significant threat to cybersecurity, necessitating effective prediction methods to identify them among numerous URLs. This paper introduces a novel approach that leverages machine learning techniques to enhance the accuracy of malicious URL prediction. The study explores various combinations of training methods and classification techniques to create a prediction model. The proposed system achieves an impressive accuracy rate of 100% using Random Forest and XG Boost technology.

## 1. INTRODUCTION

In today's interconnected digital world, the internet serves as both a conduit for seamless communication and a battleground where cyber threats constantly lurk. Among these threats, malicious Uniform Resource Locators (URLs) represent a significant and ever-evolving challenge to online security. Identifying and mitigating these malicious URLs is imperative to safeguarding individuals, organizations, and systems from the potentially devastating consequences of cyber attacks.

The motivation behind this research stems from the urgent need for effective methods to combat the proliferation of malicious URLs. Traditional approaches to URL filtering and blacklisting have proven inadequate in the face of adversaries who continuously refine their tactics to evade detection. Machine learning offers a promising avenue for enhancing URL classification accuracy by leveraging vast datasets and sophisticated algorithms to discern subtle patterns indicative of malicious intent.

The problem at hand is the accurate identification of malicious URLs within the vast expanse of internet traffic. Manual inspection and static blacklisting methods are no longer sufficient to keep

pace with the dynamic nature of cyber threats. Automated systems capable of distinguishing between benign and malicious URLs in real-time are essential for maintaining robust cybersecurity defenses.

Therefore, the objective of this research is to design and implement a novel approach for predicting and identifying malicious URLs using machine learning techniques. By exploring various combinations of training methods and classification algorithms, we aim to develop a predictive model with high accuracy and efficiency. Such a model has the potential to bolster cybersecurity measures and mitigate the risks posed by malicious URLs in an increasingly interconnected digital landscape.

## **2.LITERATURE SURVEY**

### **1. Title:** Malicious URL Detection using ML

**Author:** Mrs. Teena Varma<sup>1</sup>, Pratik Zinjad<sup>2</sup>, Shreeniket Vast<sup>3</sup>, Idris Vohra<sup>4</sup>, A.Hannan Sunsara<sup>5</sup>

**Year:** 2020

#### **Introduction**

The detection of malicious URLs is imperative in combating cyber threats and ensuring online safety. Malicious URLs can lead to various cybercrimes, including phishing attacks, malware distribution, and data breaches. Therefore, researchers have been exploring different techniques to identify and mitigate these threats effectively.

#### **Methodology:**

One common approach in malicious URL detection is the utilization of machine learning algorithms. These algorithms analyze features extracted from URLs and classify them as either malicious or benign. One such methodology involves data slicing, where datasets containing both malicious and non-malicious URLs are divided into training and testing sets. Machine learning models are trained on the training set using techniques such as repeated cross-validation to optimize accuracy and mitigate overfitting. Random Forest Algorithm, a supervised machine learning technique, has gained popularity in this domain due to its effectiveness in classification tasks.

#### **Findings**

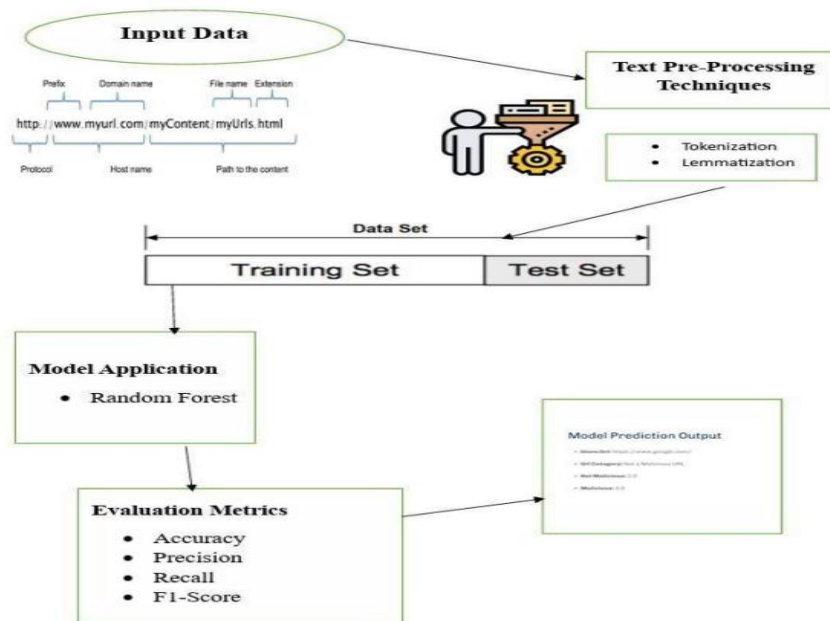
Studies utilizing machine learning algorithms for malicious URL detection have reported promising results. These algorithms demonstrate high accuracy in distinguishing between malicious and benign URLs, thereby enhancing cybersecurity measures. By leveraging features extracted from URLs and training models on diverse datasets, researchers have achieved robust detection

capabilities. Moreover, the implementation of ensemble-based methods like Random Forest Algorithm has further improved detection accuracy and resilience against adversarial attacks.

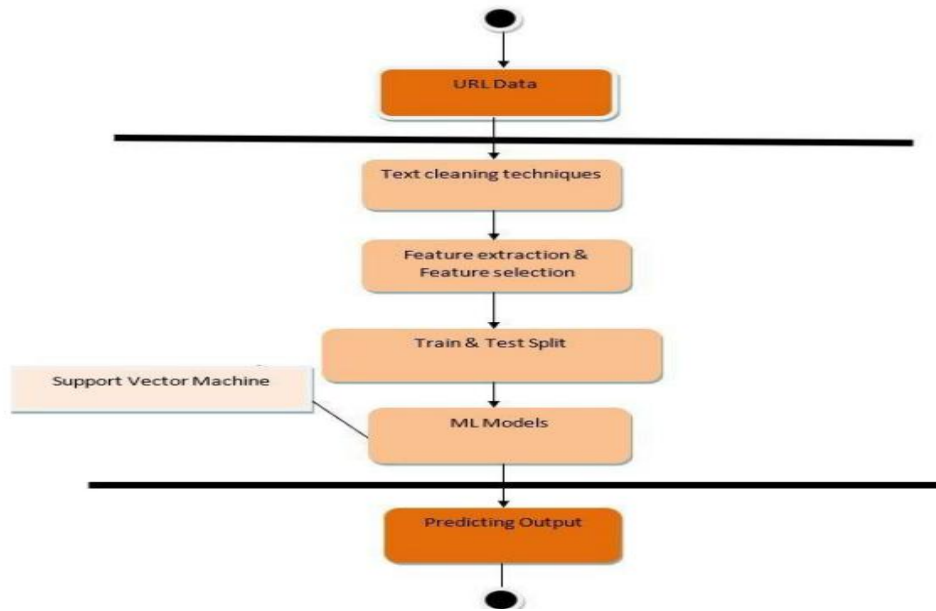
### 3. SYSTEM DESIGN

#### 3.1. Proposed system Architecture

System architecture is a conceptual model that describes the structure and behavior of multiple components and subsystems.



#### ACTIVITY DIAGRAM

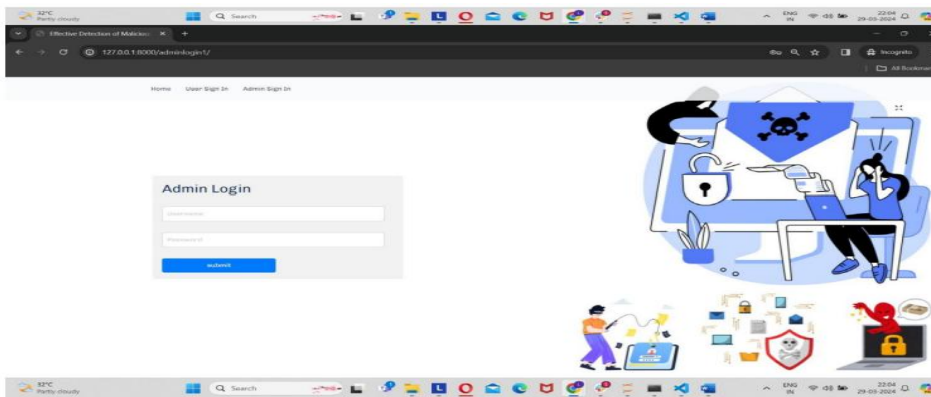
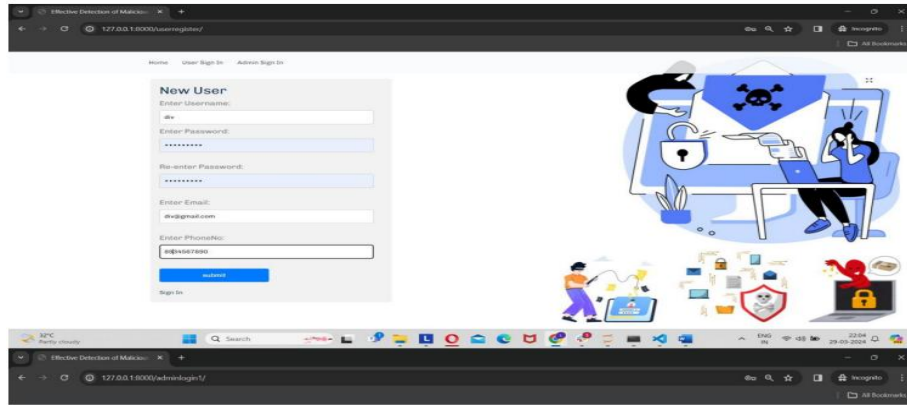


## 4. OUTPUT SCREENS

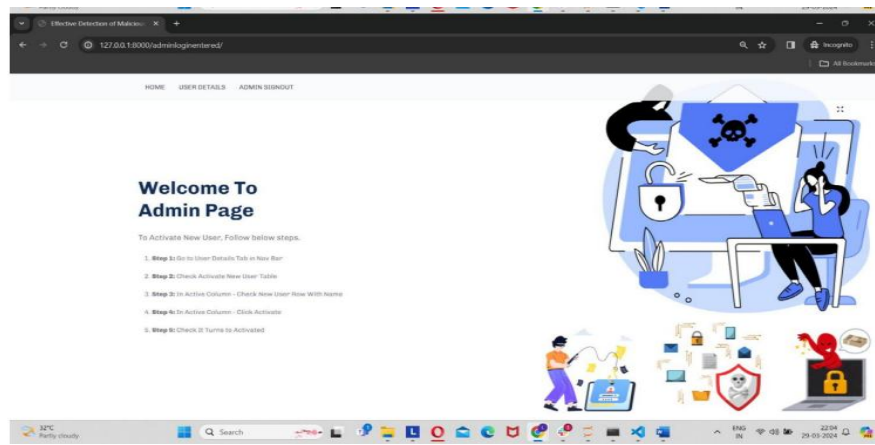
### UI Screenshots:



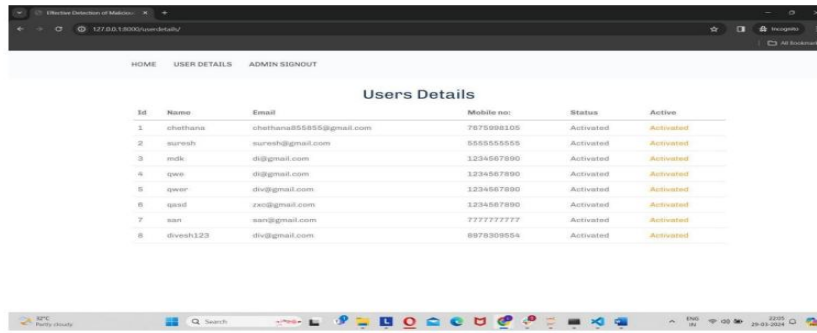
home page



registration



Admin login

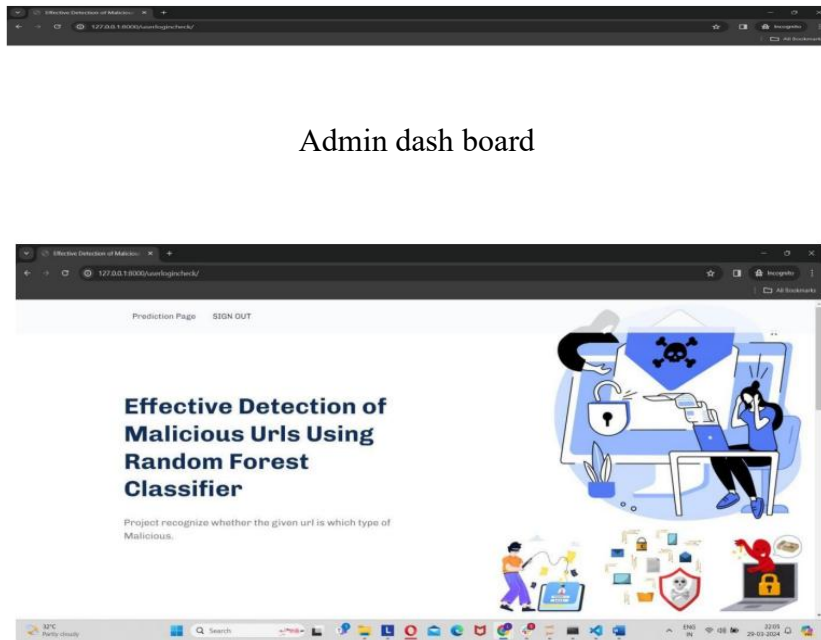


The screenshot shows a web browser window with the URL 177.88.1.83X0/admin/details/. The page has a navigation bar with 'HOME', 'USER DETAILS', and 'ADMIN SIGNOUT'. The main content is a table titled 'Users Details' with the following data:

SId	Name	Email	Mobile no:	Status	Active
1	chetana	chetana855855@gmail.com	7875998105	Activated	Activated
2	suresh	suresh@gmail.com	5555555555	Activated	Activated
3	mdk	dk@gmail.com	1234567890	Activated	Activated
4	qwe	dk@gmail.com	1234567890	Activated	Activated
5	qwer	dk@gmail.com	1234567890	Activated	Activated
6	qwe	dk@gmail.com	1234567890	Activated	Activated
7	san	san@gmail.com	7777777777	Activated	Activated
8	divesh123	dk@gmail.com	8978309554	Activated	Activated

SS 8.4 - Admin Dashboard

Admin dash board



The screenshot shows a web browser window with the URL 177.88.1.83X0/admin/predict/. The page has a navigation bar with 'Prediction Page' and 'SIGN OUT'. The main content features a large illustration of a person at a computer with a shield and a skull-and-crossbones symbol, representing security and malicious URLs. The text on the page reads:

## Effective Detection of Malicious Urls Using Random Forest Classifier

Project recognize whether the given url is which type of Malicious.

Prediction page

## 5. CONCLUSION

In conclusion, the development of effective methods for predicting malicious URLs is paramount in combating cyber threats. This study has introduced a novel approach that harnesses the power of machine learning techniques such as random forest and XGBoost to enhance prediction accuracy of 100%. Through experimentation with different combinations of training methods and classification techniques, we have constructed a robust prediction model capable of identifying malicious URLs amidst a vast pool of URLs.

Our findings demonstrate the potential of machine learning in bolstering cybersecurity efforts, offering a proactive defense mechanism against evolving threats in the digital landscape. By leveraging the wealth of data available, our approach not only improves accuracy but also enhances the efficiency of malicious URL detection, enabling swift response and mitigation strategies.

However, it is essential to acknowledge the dynamic nature of cyber threats, necessitating ongoing refinement and adaptation of prediction models to stay ahead of adversaries. Future research should focus on further optimizing model performance, exploring new features, and integrating real-time monitoring capabilities to address emerging challenges effectively.

Overall, this study contributes to the advancement of cybersecurity by providing a practical framework for predicting malicious URLs, thereby fortifying the resilience of digital infrastructures and safeguarding against potential breaches and attacks. Through continued innovation and collaboration, we can collectively work towards a safer and more secure online environment.

## 6. REFERENCES

- [1] S. He, B. Li, H. Peng, J. Xin and E. Zhang (2021), "An Effective Cost-Sensitive XG Boost Method for Malicious URLs Detection in Imbalanced Dataset," in IEEE Access, vol. 9, pp.93089-93096.
- [2] Cho Do Xuan, Hoa Dinh Nguyen, Tisenko Victor Nikolaevich (2020), "Malicious URL Detection based on Machine Learning," (IJACSA) International Journal of Advanced Computer Science and Applications", vol. 11.
- [3] Tung, Suet & Wong, Ka & Kuzminykh, Ievgeniia & Bakhshi, Taimur & Ghita, B.V.. (2022). "Using a Machine Learning Model for Malicious URL Type Detection," in ResearchGate.

- [4] Ms. Sophiya Shikalgar, Dr. S. D. Sawarkar, Mrs. Swati Narwane, "Detection of URL based Phishing Attacks using Machine Learning", Volume 08, Issue 11 (November 2019) , IJERTV8IS110269
- [5] Deepa Mary Varghese, Sreelakshmi N R, ICCIDT – 2022 (Volume 10 – Issue 04), "Phishing Website Detection using Machine Learning Techniques and CNN", IJERTCONV10IS04028.
- [6] Jason Hong, "The state of phishing attacks", Communications of the ACM, vol. 55, no. 1, pp. 74-81, 2012.
- [7] Hung Le, Quang Pham, Doyen Sahoo and CH Hoi Steven, "URLN et: Learning a URL representation with deep learning for malicious URL detection", arXiv preprint, 2018.
- [8] Doyen Sahoo, Chenghao Liu and CH Hoi Steven, "Malicious URL detection using machine learning: A survey", arXiv preprint, 2017.
- [9] R. B. Basnet, A. H. Sung, "Mining web to detect phishing URLs", Proceedings of the international conference on Machine learning and Applications, vol. 1, pp. 568-573, Dec 2012.
- [10] Alam M.S., Vuong S.T. Random Forest Classification for Detecting Android Malware; Proceedings of the 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing; Washington, DC, USA. 20–23 August 2013; pp. 663–669