IJIRSET
International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 5, May 2023

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.423

9940 572 462 | 6381 907 438 | ijirset@gmail.com | www.ijirset.com

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2023 ||

| DOI:10.15680/IJIRSET.2023.1205260 |

Phishing URL Detection Using Machine Learning

Suraj Katkar, Ajay Mahake, Yash Khandare, Dnyaneshwar Bidve

Department of Computer Engineering, Dr.D.Y. Patil School of Engineering, academy Ambi, Pune, India

ABSTRACT: Phishing attacks continue to pose a major threat for computer system defenders, often forming the first step in a multi-stage attack. There have been great strides made in phishing detection; however, some phishing emails appear to pass through filters by making simple structural and semantic changes to the messages. We tackle this problem through the use of a machine learning classifier operating on a large corpus of phishing and legitimate emails. We design SAFEPC (Semi-Automated Feature generation for Phish Classification), a system to extract features, elevating some to higher level features, that are meant to defeat common phishing email detection strategies. To evaluate SAFE-PC, we collect a large corpus of phishing emails from the central IT organization at a tier-1 university. The execution of SAFE-PC on the dataset exposes hitherto unknown insights on phishing campaigns directed at university users. SAFE-PC detects more than 70a state-of-the-art email filtering tool. It also out performs Spam Assassin, a commonly used email filtering tool.

KEYWORDS: Phishing Detection, Phishing Attack, Support Vector Machine Algorithm, Machine Learning,

I. INTRODUCTION

Phishing is a fraudulent technique that uses social and technological tricks to steal customer identification and financial credentials. Social media systems use spoofed e-mails from legitimate companies and agencies to enable users to use fake websites to divulge financial details like usernames and passwords. Hackers install malicious software on computers to steal credentials, often using systems to intercept username and passwords of consumers' online accounts.

Phishers use multiple methods, including email, Uniform Resource Locators (URL), instant messages, forum postings, telephone calls, and text messages to steal user information. The structure of phishing content is similar to the original content and trick users to access the content in order to obtain their sensitive data. The primary objective of phishing is to gain certain personal information for financial gain or use of identity theft. Phishing attacks are causing severe economic damage around the world. Moreover, Most phishing attacks target financial/payment institutions and webmail, according to the
Anti-Phishing Working Group (APWG) latest Phishing pattern studies.

II. RELATED WORK

1. Real Time Detection of Phishing Websites - This paper proposes a detection technique of phishing websites based on checking Uniform Resources Locators (URLs) of web pages. The proposed solution is able to distinguish between the legitimate web page and fake web page by checking the Uniform Resources Locators (URLs) of suspected web pages. URLs are inspected based on particular characteristics to check the phishing web pages.

2. Detection of phishing attacks - Phishing is a form of cybercrime where an attacker imitates a real person / Institution by promoting them as an official person or entity through e-mail or other communication mediums. In this type of cyber-attack, the attacker sends malicious links or attachments through phishing e-mails that can perform various functions, in cladding capturing the login credentials or account information of the victim. These e-mails harm victims because of money loss and identity thief.

3.Tools for Investigating the Phishing Attacks Dynamics Vyacheslav Lyashenko, Oleg Kobylin, MykytaMinenko. We are considering the possibility of using new tools for analysing phishing attacks. As such tools, the methods of chaos theory and the ideology of wavelet coherence are used. The use of such analysis tools makes it possible to investigate

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2023 ||

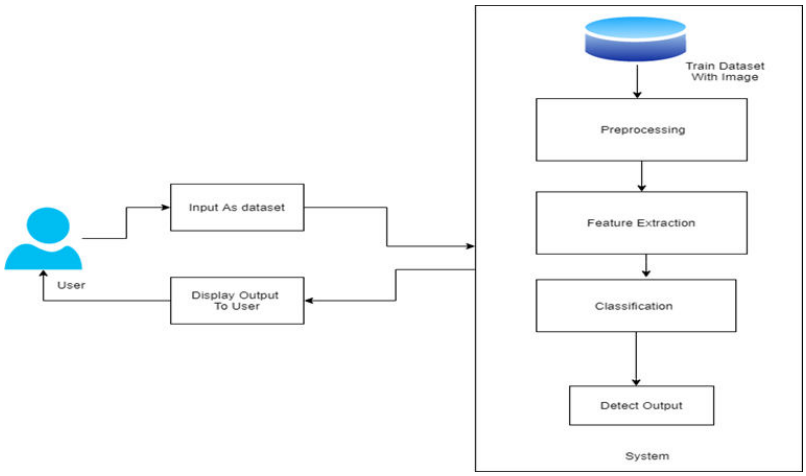
| DOI:10.15680/IJIRSET.2023.1205260 |

the peculiarities of the phishing attacks occurrence, as well as methods for their identification effectiveness. This allows you to expand the scope of the analysis of phishing attacks. For analysis, we use real data about phishing attacks

III. METHODOLOGY

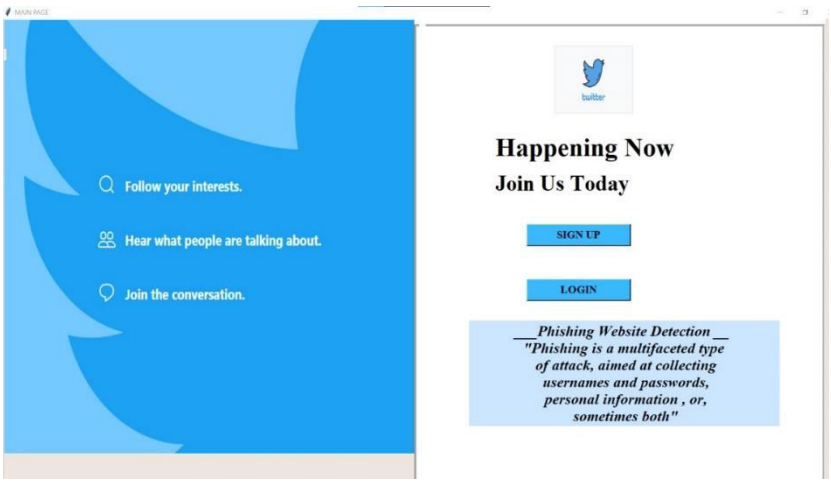
Support Vector Machine(SVM) is a supervised machine learning algorithm used for both classification and regression. Though we say regression problems as well its best suited for classification. The objective of SVM algorithm is to find a hyperplane in an N-dimensional space that distinctly classifies the data points. The dimension of the hyperplane depends upon the number of features. If the number of input features is two, then the hyperplane is just a line. If the number of input features is three, then the hyperplane becomes a 2-D plane. It becomes difficult to imagine when the number of features exceeds three.

Import the dataset. Explore the data to figure out what they look like. Pre-process the data. Split the data into attributes and labels. Divide the data into training and testing sets. Train the SVM algorithm. Make some predictions.



SystemArchitecture

IV. EXPERIMENTAL RESULTS



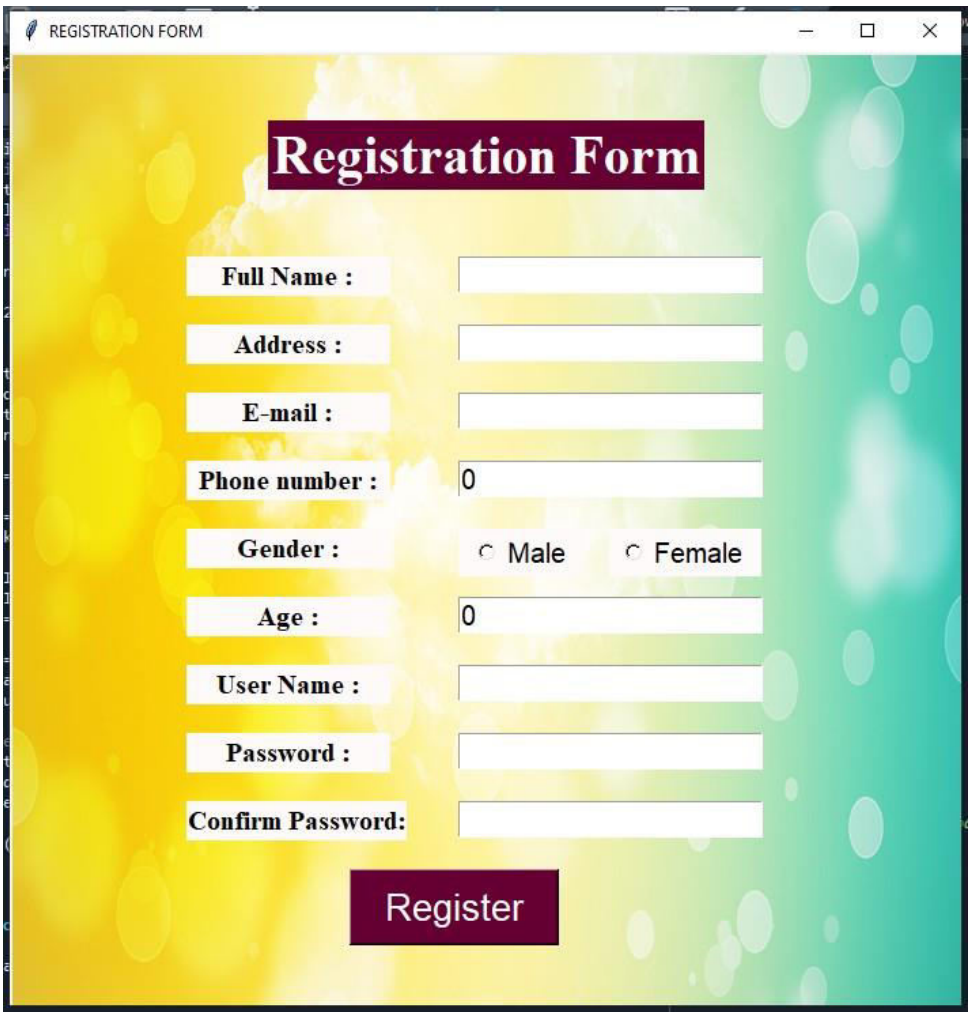
Main page

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2023 ||

| DOI:10.15680/IJIRSET.2023.1205260 |

Login Page

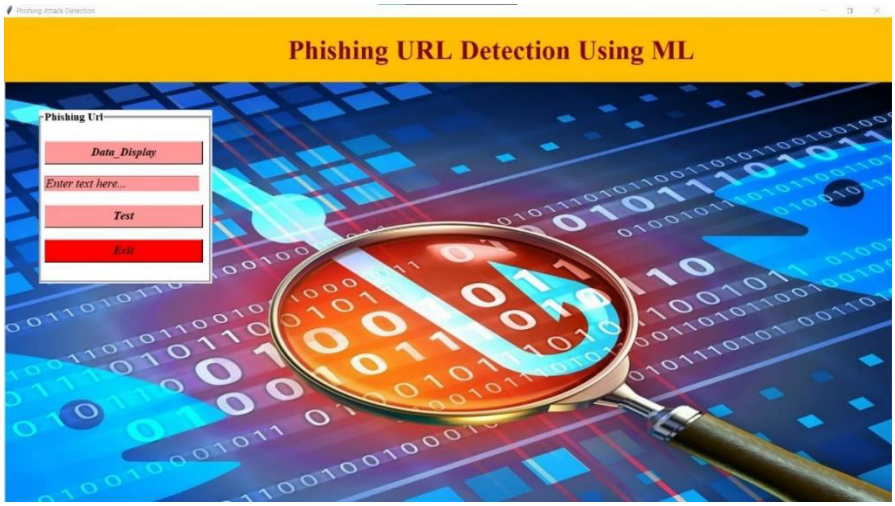


Registration Page

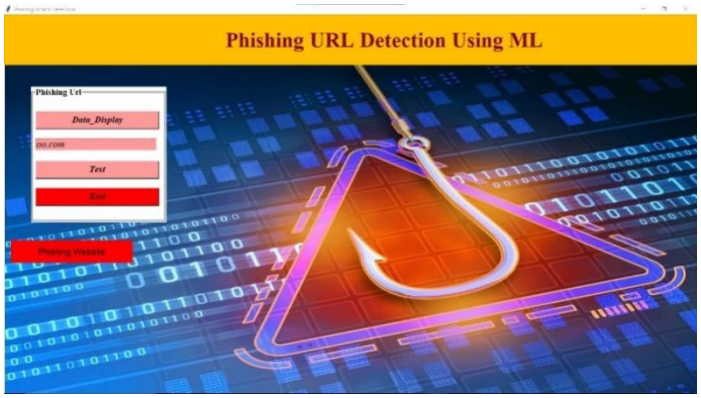
|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2023 ||

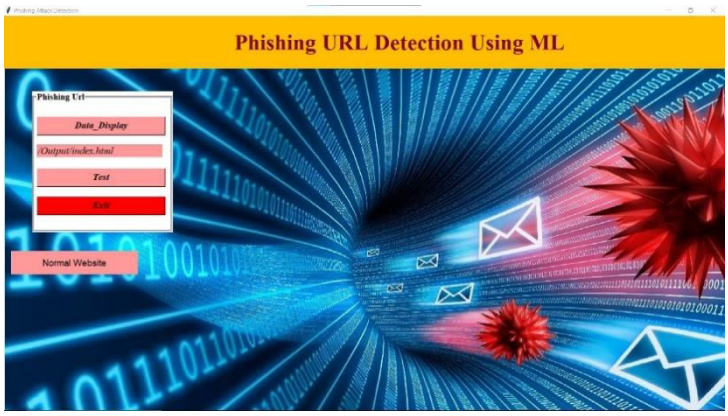
| DOI:10.15680/IJIRSET.2023.1205260 |



Testing Page



Phishing Website Result



Normal Website Result

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2023 ||

| DOI:10.15680/IJIRSET.2023.1205260 |

V. CONCLUSION

Phishing has become a serious network security problem, causing financial loss to both consumers and e-commerce companies. In this paper we've discussed about implemented system, using linkGuard algorithm.We have applied our algorithm on many websites and found that it successfully detect the website is malicious or not.

REFERENCES

1. "WC-PAD: Web Crawling based Phishing Attack Detection" Nathezhtha.T, Sangeetha.D,Vaidehi.V
2. "Detection of Phishing Attacks with Machine Learning Techniques in Cognitive Security Architecture" Ivan Ortiz-Garces, Roberto O. Andrade, and Maria Cazares.
3. "A Methodical Overview on Phishing Detection along with an OrganizedWay to Construct an Anti-Phishing Framework" SrushtiPatil, SudhirDhage.
4. "A survey of the QR code phishing: the current attacks and countermeasures" Kelvin S. C. Yong, Kang LengChiew and Choon Lin Tan.
5. "Fuzzy Rough Set Feature Selection to Enhance Phishing Attack Detection" MahdiehZabihimayvan and Derek Doran.
6. N. Goel, A. Sharma, and S. Goswami, "A way to secure a qr code: Sqr," in 2017 International Conference on Computing, Communication and Automation (ICCCA). IEEE, 2017, pp. 494–497.
7. V. Mavroeidis and M. Nicho, "Quick Response Code Secure: A Cryptographically Secure Anti-Phishing Tool for QR Code Attacks," in International Conference on Mathematical Methods, Models, and Architectures for Computer Network Security. Springer, 2017, pp. 313– 324.
8. K. Krombholz, P. Fr ¨uhwirt, T. Rieder, I. Kapsalis, J. Ullrich, and E. Weippl, "QR Code Security–How Secure and Usable Apps Can Protect Users Against Malicious QR Codes," in 2015 10th International Conference on Availability, Reliability and Security. IEEE, 2015, pp. 230–237
9. Denso Wave, "QR Code development story," 2019, [Accessed: 28-Mar2019]. [Online]. Available: https://www.denso-wave.com/en/technology/ vol1.html
10. A. Dabrowski, K. Krombholz, J. Ullrich, and E. R. Weippl, "QR inception: Barcode-in-barcode attacks," in Proceedings of the 4th ACM Workshop on Security and Privacy in Smartphones Mobile Devices. ACM, 2014, pp. 3–10

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.423

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details