IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Intelligent Methods For Accurately Detecting Phishing Websites

[1]Dr. R. S. Khule, [2]Shraddha Jadhav, [3]Nikita Sanap, [4]Shital Satote, [5]Sakshi Ugale

[1]Professor, [2]Student, [3]Student, [4]Student, [4]Student, [5]Student

[1,2,3,4,5]Information Technology Department, Matoshri College of Engineering and Research Centre, Nashik, INDIA

Abstract: Phishing attacks continue to be a major concern in the digital world, always changing and taking advantage of both human vulnerability and technological flaws, thereby putting confidential information at risk. To fight against such threats, scientists along with cybersecurity experts have started resorting more often to machine learning (ML) methods which enable automation of detection and mitigation efforts against phishing. Among these methods is Gradient Boosting Classifier which has shown excellent results as it achieved 97.4% accuracy rate and 97.7% F1 score after being trained together with other ML algorithms. This research goes deep into analyzing different attributes of URLs aiming at identifying features that can effectively differentiate legitimate links from malicious ones. These attributes include structural components like URL length, domain age or presence of subdomains as well as lexical elements i.e., text content and obfuscation techniques commonly used during phishing attacks among others. Furthermore, semantic aspects take into consideration contextual understanding of URLs through examination web page contents, SSL certificates plus behavioral patterns associated with bad domains. By utilizing all these diverse characteristics, it becomes possible for machine learning algorithms to finely tune themselves so that they can accurately detect phishing URLs with high levels of precision and recall rates. However, the research also recognizes that there are difficulties and limitations when it comes to identifying phishing URLs with machine learning. These include imbalances in datasets, complex feature engineering, and use of adversarial evasion techniques by advanced attackers. Still, this doesn't mean that Machine Learning is not effective against phishing attacks; in fact, its effectiveness is still significant especially shown by performances of Gradient Boosting Classifier**.**

Index Terms **- Detection of Phishing URLs, Cybersecurity Measures, Extracting Features, Supervised Learning Techniques, Gradient Boosting, Recognizing Patterns.**

I. INTRODUCTION

Phishing attacks persist as a significant and ever-changing cybersecurity menace which employs tricks to lure users into revealing sensitive information or participating in harmful actions. Traditional approaches for detecting phishing often fail to keep pace with the dynamic strategies employed by attackers who never stop evolving them. This has led many experts in this field, including both scientists and practitioners, to rely heavily on machine learning (ML) techniques for developing automated systems capable of recognizing and neutralizing different types of phishing threats.

One necessary part of this strategy is to use machine learning (ML) algorithms for scrutinizing Uniform Resource Locators (URLs). URLs, the web addresses used to access online resources, are entry points for users' interaction with fake content or websites pretending to be genuine organizations hence their significance in phishing attacks. Different aspects such as construction, language and setting among others can be examined by ML models to identify patterns that show phishing behavior through URLs. From a statistical point of view, design features like length of URL; subdomains presence as well age of domain name provides useful insights about link credibility. Lexical properties entail looking closer at words used in URLs for signs related to

phishing or techniques used to hide them. Moreover, semantic characteristics consider webpage contents, SSL certificates as well abnormal behaviors associated with malicious domains thus going beyond context and meaning inherent in uniform resource locators themselves alone. These various inputs allow machine learning systems to appreciate the complex nature of phishing attacks thus differentiating between legit and harmful links accurately often.

Cybersecurity is constantly threatened by phishing attacks hence the need to automate detection through machine learning (ML). To detect behavior patterns associated with phishing, models of machine learning (ML) investigate different parts of Uniform Resource Locators (URLs) including structure, language and context. Although it is an approach that requires many different disciplines, there are some difficulties that this approach faces which include imbalanced datasets; intricate feature engineering and privacy issues with collecting or analyzing data on URLs in terms of their usage for other purposes such as advertising. All these challenges can only be overcome if researchers work together alongside policy makers plus those involved with industry who have got much knowledge in this area. By so doing their combined efforts will help advance cyber security powered by ML thus safeguarding various online spaces.

II. RELATED WORK

The authors of this study [1] are discussing defensive strategies against phishing attacks, which are a common type of cyber threat. The creation of DeepPhish, a system for creating URLs for phishing scams based on deep neural networks, has made it necessary to develop more reliable detection methods. Therefore, to solve this problem, they propose PhishHaven – an ensemble machine learning-based detection system that can recognize both AI-generated and human-crafted phishing URLs among other things. This research is important because until now no one has tried to detect phishing campaigns launched by machines and people at the same time. PhishHaven uses lexical analysis for feature extraction and employs several innovative techniques such as URL HTML Encoding which help improve its ability to detect small-sized URLs not solved by any other method available today necessarily designed for large ones only. Moreover, within PhishHaven authors realized URL Hit approach together with unbiased voting mechanism aimed at accurate classification with mitigation of misclassification cases too. The practicality of PhishHaven in fighting phishing threats is demonstrated when it is applied for real-time detection with the help of multi-threading. In this regard, the authors also offer some theoretical reflections on their solution showing how it can always find small and future AI-generated URL based phishing accurately using specific lexical features with 100% precision. The study further goes on to validate empirically that PhishHaven effectively works against different types of cyber threats by testing against one hundred thousand samples drawn from an industry standard benchmark dataset containing both malicious and benign web addresses, achieving a remarkable success rate of 98%. This achievement outperforms any existing systems which rely solely on lexical analysis techniques to detect human-crafted phishing URLs thus highlighting its efficiency over other approaches designed for handling evolving cyber security challenges. What makes this paper unique among many others dealing with anti-phishing methods lies not only in its innovative approach but also in providing new ideas as well as tools for use within wider areas pertaining computer safety research and defense mechanisms against attacks from external actors who may want access private networks or steal sensitive information.

In this research paper [2], the authors address the problem of bad activities on Twitter such as sending spam, phishing, and malware through URLs shared in tweets. They evaluate existing detection methods critically and point out that these methods have not been effective enough at identifying and mitigating such threats. Traditional approaches rely on either account features or relational aspects in the Twitter graph, but they often fail to work because they can be deceived by fabricated features or because they require too many resources for processing. Similarly, suspicious URL detection techniques which use different kinds of features like lexical analysis and dynamic behavior assessment can be evaded using time-based evasion or crawler evasion methods. To overcome these challenges, WARNINGBIRD is introduced as a new system that is specifically designed for detecting suspicious URLs within the Twitter ecosystem. WARNINGBIRD detects correlations among URL redirect chains from multiple tweets. It works by investigating the tendency of attackers to reuse limited resources. The system identifies frequently shared URLs in these chains, recognizes suspicious patterns and assesses their risk level. To effectively detect and classify suspicious URLs, the authors of this paper use a statistical classifier built on large Twitter dataset collected from its public timeline. They show that WARNINGBIRD can accurately find malicious URLs while also being suitable for monitoring real-time Twitter streams. This research not only provides an effective way of dealing with internet threats but also gives a useful solution based on specific features of the Twitter platform.

In this research [3], the author explores the intricate field of user vulnerability to phishing attacks through conducting a scientific study and meta-analysis of previous findings. The main objective is to provide a holistic view about what influences susceptibility towards phishing attacks basing on various studies but with emphasis on age and sex differences. The results are mixed, showing different outcomes in different experiments. Though it might be assumed that older people are more likely to fall for these scams, over half of all research examined found no significant correlation between age and performance among users. There are also variations concerning gender; certain investigations pointed out females' higher vulnerability while others showed no difference at all. The Author's meta-analysis, which was conducted after careful analysis, has revealed many significant findings. The first finding is that age does have a role in determining susceptibility to phishing scams as some assumptions had it before now were wrong. Again, the outcomes show that ladies are truly more prone than boys thereby indicating an intervention point or public enlightenment campaign may be necessary. Ultimately, this study demonstrates how effective user education can be at improving detection skills thus presenting a preventive measure against phishing attacks. This detailed investigation contributes greatly towards appreciating vulnerability when it comes to such scams while also giving useful hints for further studies and practices in cyber security.

III. PROPOSED WORK

Precisely what we did for our phishing URL detection system so that it would be strong enough was a step-by-step process. Firstly, we began off by loading the dataset which consists of many URL features and their corresponding labels carefully. This acted as a base for the rest of our analysis and training models later. After that, we waded into exploratory data analysis (EDA) to further understand the complexities inherent within the dataset itself. By examining how the information was distributed, structured, and patterned, different insights were gained into what could be done during modeling stage. Once all these stages had been completed — getting used with data followed by EDA — it became necessary that some visualizations should be created so as better comprehend various aspects contained within this dataset type through diverse range visualization techniques employed. We made different types of plots, graphs and charts which helped us understand how each feature related with others affecting target variable more clearly hence guiding next steps in model building phase accordingly. Informed by this understanding of the dataset, we divided it into training set and testing sets since now we can see different components or variables interact among themselves better also this will enable evaluation after model fitting process.

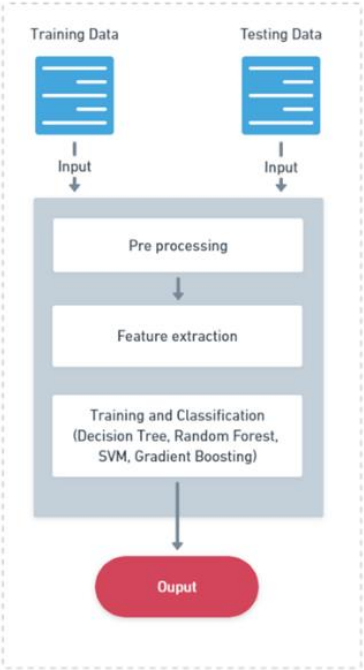

Figure 1 - Architecture of Proposed System

At the heart of our approach was the strict model training and evaluation process that involved utilizing machine learning algorithms across a wide range to develop predictive models from the training data. The Gradient Boosting Classifier proved to be the best among these algorithms as it recorded unmatched accuracy levels and F1 score metrics. After making comparisons thoughtfully, we discovered that this classifier was more effective than any other one thus choosing it as a model for real-life applications. To implement this solution, we wrapped up our trained model in addition to embedding it on a user-friendly interface (UI) created using Streamlit so that end-users can have an easy-to-use platform for detecting phishing URLs. We also make sure that before giving predictions, all URL features are covered by this system hence making it more strong and accurate in identifying potentially malicious URLs which helps beef up cyber defense posture both for organizations and individuals at large.

IV. RESULT AND ANALYSIS

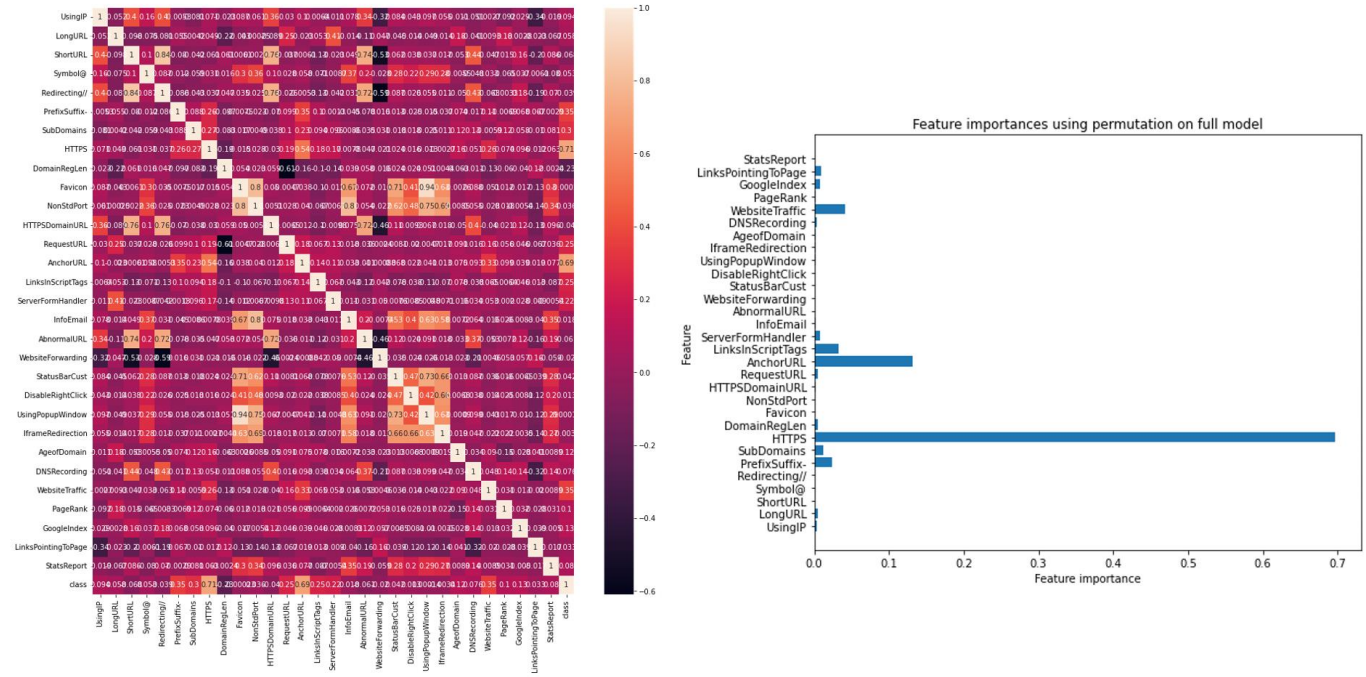

Figure 2 - Result (Heatmap and Feature Importance)

The results are promising for the Gradient Boosting Classifier model, which appears to be good at detecting phishing URLs because it has achieved high accuracy, F1 score, recall and precision on training as well as test data. Still, there is a potential for improvement such as decreasing false positives (i.e., increasing precision) on the test data. The model could be refined further by analyzing feature importance or trying out different hyperparameters for better performance.

Table 2. Accuracy of various model

No.	ML Model	Accuracy	F1 Score	Recall	Precision
1	Gradient Boosting Classifier	97.4%	97.7%	99.4%	98.6%
2	Random Forest	96.7%	99.3%	99.3%	99.0%
3	Support Vector Machine	96.4%	96.8%	98.0%	96.5%
4	Naive Bayes Classifier	60.5%	45.4%	29.2%	99.7%

According to this information, none of the machine learning models have high accuracy rates like those shown by Gradient Boosting Classifier in terms of accuracy, F1 score, recall and precision. In all the metrics used here Gradient Boosting Classifier surpasses all other models indicating its ability to accurately identify phishing

URLs more frequently than any other method. This may be because it can keep refining itself through successive attempts at difficult problems thus making it very powerful in this particular case.

V. CONCLUSION

We have made and tested the Gradient Boosting Classifier model for Phishing-URL detection. It gave good results in different evaluation metrics (accuracy, F1 score, recall and precision) on various train-test splits. In fact, this model can be used as a very effective tool for identifying legitimate URLs from those that are phishing based due to its 98.9% accuracy on training data and 97.4% accuracy on test data. Another thing to note about the F1 score which considers both precision & recall is that it achieved 99.0% (training data) and 97.7% (test data), thus striking a good balance between correctly detecting phishing URLs while minimizing false positives here too. The recall score also deserves special mention since it indicates how many of all actual phishing URLs were caught by our system; thus, showing higher figures such as 99.4% (training data) or even 98.9% (test data). Nonetheless, there were some misclassifications made by the algorithm according to these measures' values: precision scores were not perfect – they equaled 98%. More exactly speaking about misclassification types revealed during testing phase one can mention false positives especially being detected at rate below expectations from greater than ninety-six percent level up until reaching ninety-eight-point six percent marks – so this area needs further improvement if reliability should be increased for accurate identification of valid URLs by models like ours.

In the identification of phishing URLs, the Gradient Boosting Classifier model shows potential by obtaining good results that are characterized by high accuracy levels in terms of F1 score, recall, and precision. The efficiency of this model can be seen through all these measures combined because it is able to differentiate between fake and genuine web addresses while also ensuring that there is enough detection rate for false positives on phishing sites. There are a few things which could be done to enhance its precision further and make it more adaptable in detecting real-life phishing cases such as parameter adjustment or looking into other methods like assessing feature importance among others.

REFERENCES

[1] M. Sameen, K. Han and S. O. Hwang, "PhishHaven—An Efficient Real-Time AI Phishing URLs Detection System," in IEEE Access, vol. 8, pp. 83425-83443, 2020, doi: 10.1109/ACCESS.2020.2991403

[2] S. Lee and J. Kim, "WarningBird: A Near Real-Time Detection System for Suspicious URLs in Twitter Stream," in IEEE Transactions on Dependable and Secure Computing, vol. 10, no. 3, pp. 183-195, May-June 2013, doi: 10.1109/TDSC.2013.3

[3] S. Baki and R. M. Verma, "Sixteen Years of Phishing User Studies: What Have We Learned?," in IEEE Transactions on Dependable and Secure Computing, vol. 20, no. 2, pp. 1200-1212, 1 March-April 2023, doi: 10.1109/TDSC.2022.3151103.

[4] L. Wu, X. Du and J. Wu, "Effective Defense Schemes for Phishing Attacks on Mobile Computing Platforms," in IEEE Transactions on Vehicular Technology, vol. 65, no. 8, pp. 6678-6691, Aug. 2016, doi: 10.1109/TVT.2015.2472993

[5] R. R. Rout, G. Lingam and D. V. L. N. Somayajulu, "Detection of Malicious Social Bots Using Learning Automata with URL Features in Twitter Network," in IEEE Transactions on Computational Social Systems, vol. 7, no. 4, pp. 1004-1018, Aug. 2020, doi: 10.1109/TCSS.2020.2992223.

[6] E. Nowroozi, Abhishek, M. Mohammadi and M. Conti, "An Adversarial Attack Analysis on Malicious Advertisement URL Detection Framework," in IEEE Transactions on Network and Service Management, vol. 20, no. 2, pp. 1332-1344, June 2023, doi: 10.1109/TNSM.2022.3225217.

[7] Y. Sönmez, T. Tuncer, H. Gökal and E. Avcı, "Phishing web sites features classification based on extreme learning machine," 2018 6th International Symposium on Digital Forensic and Security (ISDFS), Antalya, Turkey, 2018, pp. 1-5, doi: 10.1109/ISDFS.2018.8355342.

[8] M. H. Alkawaz, S. J. Steven and A. I. Hajamydeen, "Detecting Phishing Website Using Machine Learning," 2020 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA), Langkawi, Malaysia, 2020, pp. 111-114, doi: 10.1109/CSPA48992.2020.9068728.