Detection of Phishing Attacks using Machine Learning

Dr. J. Mala[1], R. Dhinesh[2], T. Jeevanandham[3], T. Thinesh[4]

[1] Assistant Professor, Department of Information Technology, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, Email: mala.it@srit.org.

[2] Student, Department of Information Technology, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, Email: dhinesh.2006@srit.org.

[3] Student, Department of Information Technology, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, Email: jeevanandham.2006@srit.org.

[4] Student, Department of Information Technology, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, Email: thinesh.2006@srit.org.

Abstract

The use of the internet in day-to-day activities has increased the risk of phishing attacks, which put user's sensitive data at serious risk. A thorough phishing detection system that uses both static and dynamic features to find possible threats is presented. A web page's legitimacy is evaluated by the system using thirty different features, which it then divides into three categories: legitimate, suspicious, and phishing. Websites are grouped into these categories using machine learning techniques. The evaluation results demonstrate how well the system works to identify malicious websites from legitimate ones while reducing false positives and false negatives. Additionally integrated is a user feedback mechanism that enables users to report doubtful URLs and help improve the system. This research advances cybersecurity by presenting a comprehensive and flexible solution for phishing detection, giving an all-encompassing view of the properties of web pages, and strengthening their resistance to advanced phishing tactics.

Keywords: Phishing Detection, Machine Learning, Feature Analysis, Web Security, User Feedback.

1. Introduction

Web security has become increasingly important due to the increasing number of people who are using the internet. Phishing attacks are a persistent threat to users because they use deceptive methods to trick people into revealing sensitive information like usernames, passwords, and financial details. These attacks may result in identity theft or other serious financial losses, among other unfavourable outcomes. This paper or article addresses the pressing need, given this landscape, for trustworthy and efficient systems to quickly detect phishing attacks.

Traditional detection methods have become less effective due to the evolving sophistication of phishing attacks, mainly because of their increased

complexity. Cultured systems that can adjust to the ever-changing nature of phishing tactics are essential as cybercriminals constantly improve their methods. This paper or article suggests a thorough strategy to improve web security through the creation and application of advanced technology that can quickly detect and reduce the risks connected to phishing attacks.

This paper not only examines the nuances of phishing attacks and their possible outcomes, but it also looks into new tactics for enhancing cybersecurity defences. The conversation explores the creation of adaptable systems that make use of cutting-edge technology to fend off new threats. This paper adds to the ongoing discussion on improving web security in the face of evolving cyber threats by offering a thorough analysis of the shortcomings of conventional detection methods and suggesting creative solutions.

Background:

Cybercriminals use phishing attacks, which are advanced psychological techniques, to trick users into disclosing personal information. These attacks take advantage of human weaknesses by employing strategies like curiosity, fear, or urgency. The ongoing development of these strategies puts traditional phishing detection techniques to the test. To strengthen cybersecurity against the ever-present and changing threat of phishing attacks, it is imperative to comprehend the dynamic nature of these attacks.

2. Related Work

In recent years, researchers have made significant steps in the domain of phishing detection, employing various techniques and algorithms to enhance accuracy and robustness.

Dhanalakshmi Ranganayakulu and Chellappan .C explored the utilization of the URLs classifier method in "Detecting Malicious URLs in E-mail – An Implementation". [1] Their method demonstrated high efficacy, achieving an accuracy of 92.8%, providing a reliable mechanism for detecting phishing URLs.

In "Machine Intelligence Based Web Page Phishing Detection," [2] Dhruv Soni and Kalpita Gadhe and Bhavya Doshi and Amitava Choudhury introduced autoML and convolution neural network (CNN) techniques, attaining an impressive accuracy of 85%. This suggests the potential of advanced neural network architectures in accurately discerning phishing attempts.

Dogukan Aksu and Zeynep Turgut and Serpil Ustebay and Aydin .M adopted a comprehensive approach in "Phishing Analysis of Websites Using Classification Techniques". [3] By employing neural networks, support vector machines, decision trees, and stacked autoencoders, they achieved an accuracy of 86%, showcasing the effectiveness of a multi-model classification strategy in identifying phishing websites.

Gunikhan Sonowal and Kuppusamy .K presented the PhiDMA model in "A phishing detection model with a multi-filter approach," [4] incorporating five layers. The model demonstrated robustness with an accuracy of 92.72%, emphasizing the importance of multi-filter strategies in phishing detection.

Jian Mao and Wenqian Tian and Pei Li and Tao Wei and Zhenkai Liang focused on Cascading Style Sheets (CSS) in "Phishing-Alarm: Robust and Efficient Phishing Detection via Page Component Similarity". [5] Their approach, centered around CSS analysis, achieved an impressive F1-score of

99%, showcasing the significance of considering the visual aspects of web pages in phishing detection.

Jiwon Hong and Taeri Kim and Jing Liu and Noseong Park and Sang-Wook Kim contributed to the field by improving the classification lexical features using blacklisted domains in "Phishing URL Detection with Lexical Features and Blacklisted Domains". [6] The F-1 method accuracy of 84% indicates the effectiveness of lexical analysis combined with domain-based blacklisting.

Muzammil Ahmed and Altamimi .A.B and Wilayat Khan and Alsaffar .M and Aakash Ahmad and Khan .Z and Abdulrahman Alreshidi proposed an approach in "PhishCatcher: Client-Side Defense Against Web Spoofing Attacks Using Machine Learning". [7] By leveraging the Random Forest algorithm, they achieved a notable accuracy of 94.5%, surpassing the performance of other machine learning algorithms.

Naresh R and Ayon Gupta and Sanghamitra Giri introduced the combined use of SVM classification & Logistic Regression in "Malicious URL Detection System Using Combined SYM and Logistic Regression Model," [8] achieving an accuracy of 98%. This suggests that combining different classifiers can lead to a more robust phishing detection system.

In a survey-based approach, Rasha Zieni and Massari . L and Calzarossa .M explored list-based, similarity-based, and machine learning-based techniques in "Phishing or Not Phishing? A Survey on the Detection of Phishing Websites". [9] The study showcased the versatility of combining various techniques, resulting in improved accuracy compared to individual methods

Routhu Srinivasa Rao and Syed Taqi Ali introduced a Heuristic method in "PhishShield: A Desktop Application to Detect Phishing Webpages through Heuristic Approach," [10] achieving an accuracy of 96.57%. The heuristic approach provides an additional layer of defense by leveraging rule-based systems.

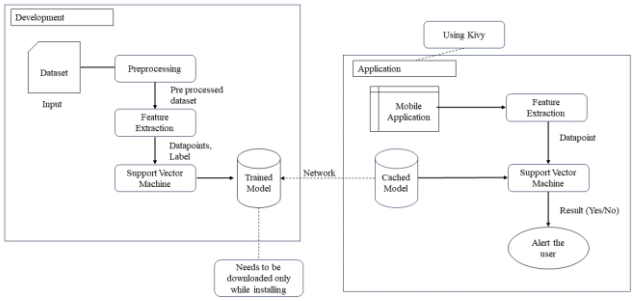
This collective body of work underscores the diverse strategies and algorithms employed in phishing detection, offering valuable insights for the development of more effective and robust systems.

3. Methodology

The Phishing Detection App uses a thorough approach to differentiate between websites that are potentially phishing and those that are not. Using a wide range of features that are taken from HTML content, domain information, and other attributes, the model takes a multidimensional approach to precise detection. These characteristics include indicators like the use of HTTPS, traits specific to a domain, anomalies in URLs, and scripting behaviours. To improve its discriminatory power, the model also evaluates factors such as age of domain, email submission forms, and redirects. The incorporation of rule-based heuristics, such as pop-up window generation and checks for right-click disabling, facilitates prompt decision-making by utilising established patterns.

The Support Vector Machine (SVM) algorithm enhances the fundamental characteristics of the model, enabling it to adjust to changing phishing strategies. The SVM model guarantees a strong and adaptable phishing detection system by improving the system's ability to identify both known and unknown threats. The integration of rule-based and machine learning methodologies highlights the robustness and precision of the model in navigating through the ever-changing terrain of cyber threats. The system is able to recognise understated and changing malicious patterns because of the design

choices, which represent a strategic combination of features to address the complex nature of phishing attacks.



1: Architecture diagram

4. Experimental Setup

Dataset Description: The phishing detection model is experimentally evaluated using a dataset that contains more than 11,000 website URLs. Every sample in the dataset includes 30 website parameters, from server behaviours to URL characteristics, along with a class label designating the phishing status (1 for legitimate, -1 for phishing). The parameters provide a comprehensive view of website attributes, encompassing a variety of features like URL length, the presence of specific symbols, HTTPS usage, domain registration length, and more.

Evaluation Metrics: The phishing detection model is evaluated using common metrics that are essential for binary classification tasks. Crucial metrics that reveal how well the model can detect phishing instances while reducing false positives are precision, recall, and F1-score.

1. Accuracy: How well the model predicts things overall.

Accuracy=TP + TN / TP + TN + FP + FN

2. Precision: The proportion of accurately predicted phishing incidents to all predicted phishing incidents.

Precision = TP / TP + FP

3. Recall (Sensitivity or True Positive Rate): The proportion of correctly predicted to actual phishing incidents.

Recall = TP / TP + FN.

4. F1 Score: A balanced metric produced by taking the harmonic mean of recall and precision.

F1score = 2 × (Recall x Precision) / (Recall + Precision).

Where,

- TP (True Positives) is the number of correctly predicted positive instances.
- TN (True Negatives) is the number of correctly predicted negative instances.
- FP (False Positives) is the number of incorrectly predicted positive instances.
- FN (False Negatives) is the number of incorrectly predicted negative instances.

These measures are frequently employed to assess how well classification models perform.

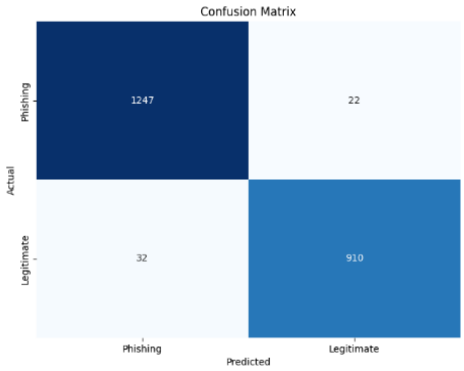
Preprocessing Steps: To guarantee optimal performance, the dataset is carefully pre-processed before model training. This includes encoding categorical variables, standardising numeric features to a scale, and handling inconsistent or missing data. The experimental setup is intended to offer a thorough evaluation of the phishing detection model by taking into account a variety of website attributes and using reliable assessment metrics to gauge the model's efficacy and generalizability.

5. Results

The evaluation results of different machine learning classifiers for phishing detection on the dataset are shown in this section. Several classifiers were used, such as SVM, Random Forest, XGBoost, KNN, Logistic Regression, Naive Bayes, and Decision Tree. Using GridSearchCV, the classifiers were optimised, and their effectiveness was evaluated in terms of accuracy, precision, recall, and F1 score.

The results of each classifier are compiled in Table 1, which also highlights the optimal parameters, accuracy, precision, recall, and F1 score.

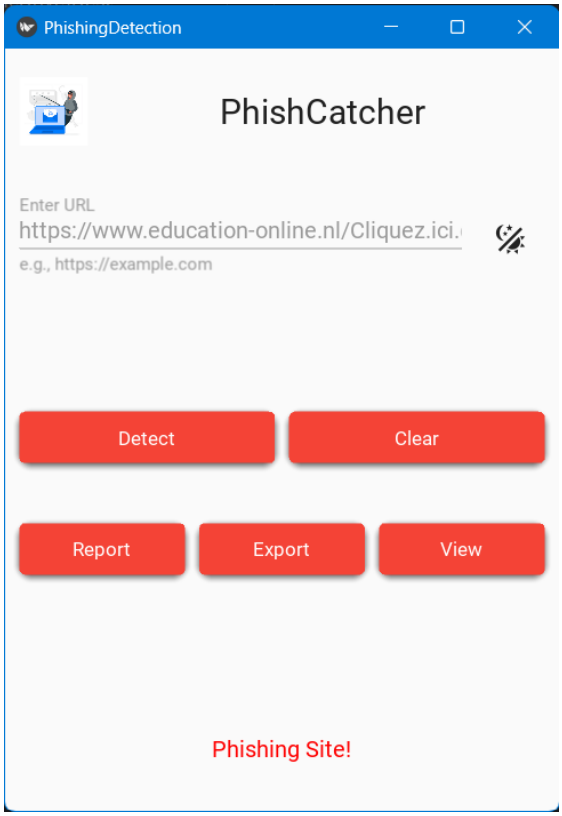
Confusion matrix:



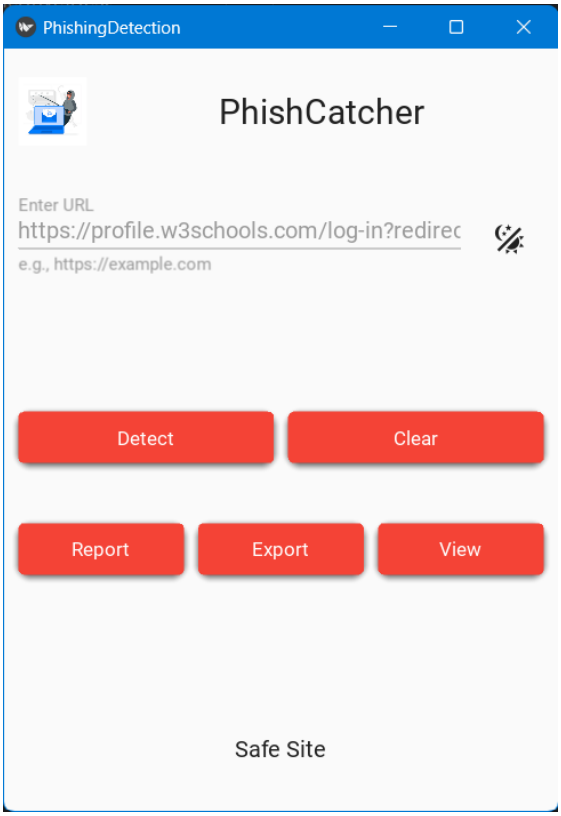
2: Confusion Matrix for SVM Algorithm

Table 1: Model Evaluation Results

Model	Best Parameters	Accuracy	Precision	Recall	F1 Score
Random Forest	bootstrap: False, max_depth: 30, min_samples_leaf: 1, min_samples_split: 5, n_estimators: 100	97.33%	97.33%	97.33%	97.69%
SVM	C: 10, gamma: 0.1, kernel: rbf	97.56%	97.56%	97.56%	97.88%
XGBoost	colsample_bytree: 0.8, learning_rate: 0.1, max_depth: 7, n_estimators: 300, subsample: 0.9	96.88%	96.88%	96.88%	97.29%
KNN	n_neighbors: 7, p: 1, weights: distance	96.97%	96.98%	96.97%	97.38%
Logistic Regression	C: 0.1, penalty: l2	92.76%	92.76%	92.76%	93.72%
Naive Bayes		58.66%	78.86%	58.66%	43.79%
Decision Tree	criterion: entropy, max_depth: None, min_samples_leaf: 1, min_samples_split: 2	96.16%	96.15%	96.16%	96.66%



3: Result for Phishing Site



4: Result for Safe Site

confusion matrix illustrating the effectiveness of the phishing detection model shows that it correctly classified 910 legitimate sites and successfully identified 1247 instances of phishing sites. Nevertheless, 22 cases of phishing were mistakenly classified as legitimate, and 32 cases of legitimate websites were mistakenly classified as phishing. The matrix offers a thorough summary of the model's advantages and disadvantages. An interactive interface with real-time classification capabilities and instructional materials can be used to improve user engagement. A feedback mechanism also promotes a collaborative approach to increasing accuracy and efficacy by enabling users to make gradual improvements to the model.

6. Discussion and Analysis

The findings show how well various classifiers work to identify phishing websites. With an accuracy of 97.56%, SVM was the most accurate, closely followed by Random Forest at 97.33%. Both models performed well according to a variety of metrics.

The best-performing classifier was the SVM model with the ideal parameters {'C': 10, 'gamma': 0.1, 'kernel': 'rbf'}, demonstrating its applicability for phishing detection in this situation. The SVM model's dependability in producing precise predictions is further supported by the metrics of precision, recall, and F1 score. These results provide insightful information about how to choose classifiers for phishing detection applications by taking into account variables like accuracy, precision, and recall according to particular needs and application priorities.

Limitations and Strengths: The model's strength is its extensive feature set, enabling it to identify subtle patterns linked to phishing websites. However, it's

important to recognise the limitations, such as possible difficulties managing phishing techniques that are constantly evolving. The representativeness and level of the training dataset may also affect the model's performance.

7. Conclusion

By offering a phishing detection model that excels in accuracy and dependability, this paper makes major contributions to the field of web security. Using SVM in conjunction with a large feature set improves the model's resistance to changing phishing techniques. The interface of the application is easy to use, which encourages wider adoption. The findings have consequences for users, organisations, and cybersecurity professionals. Because of the model's accuracy, users who browse the internet can feel even more protected. By integrating this solution, organisations can strengthen their security infrastructure and reduce the risks brought on by phishing attacks.

References

[1] Dhanalakshmi Ranganayakulu and Chellappan .C, “Detecting Malicious URLs in E-mail – An Implementation”, AASRI Procedia, pp 125-131, vol.4, 2013.

[2] Dhruv Soni and Kalpita Gadhe and Bhavya Doshi and Amitava Choudhury, “Machine Intelligence Based Web Page Phishing Detection”, International Conference on Futuristic Technologies (INCOFT), IEEE Access, 2022.

[3] Dogukan Aksu and Zeynep Turgut and Serpil Ustebay and Aydin .M, “Phishing Analysis of Websites Using Classification Techniques”, Lecture Notes in Electrical Engineering, International Telecommunications Conference , SpringerLink, pp 251–258, vol.504, 2018.

[4] Gunikhan Sonowal and Kuppusamy .K, “PhiDMA - A phishing detection model with multi-filter approach”, King Saud .J Univ. Comput. Inf. Sci, ScienceDirect, Issue 1, pp 99-112, vol.32, 2017.

[5] Jian Mao and Wenqian Tian and Pei Li and Tao Wei and Zhenkai Liang, “Phishing-Alarm: Robust and Efficient Phishing Detection via Page Component Similarity”, IEEE Access, pp 17020 – 17030, vol .5, 2017.

[6] Jiwon Hong and Taeri Kim and Jing Liu and Noseong Park and Sang-Wook Kim, “Phishing URL Detection with Lexical Features and Blacklisted Domains”, Adaptive Autonomous Secure Cyber Systems, SpringerLink, pp 253–267, 2020.

[7] Muzammil Ahmed and Altamimi .A.B and Wilayat Khan and Alsaffar .M and Aakash Ahmad and Khan .Z and Abdulrahman Alreshidi, ”PhishCatcher: Client-Side Defense Against Web Spoofing Attacks Using Machine Learning”, pp 61249-61263, IEEE Access, vol.11, 2023.

[8] Naresh R and Ayon Gupta and Sanghamitra Giri, “Malicious URL Detection System Using Combined SYM and Logistic Regression Model”, InfoSciRN: Information Architecture (Topic), International Journal of Advanced Research in Engineering and Technology, SSRN, pp. 63-73, 2020.

[9] Rasha Zieni and Massari . L and Calzarossa .M, “Phishing or Not Phishing? A Survey on the Detection of Phishing Websites”, IEEE Access, pp18499- 18519, vol. 11, 2023.

[10] Routhu Srinivasa Rao and Syed Taqi Ali, “PhishShield: A Desktop Application to Detect Phishing Webpages through Heuristic Approach”, Procedia Computer Science, ScienceDirect, vol 54, pp 147-156, 2015.