Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 6s (2025)

Automated Phishing Detection Through URL Analysis and Machine Learning

Vijaykumar[1], Basavaraj G N[2], Mohan Bangalore Anjaneyalu[3], Swetha M S[4], E G Satish[5]

[1,2,3,4]Department of Information Science, BMS Institute of Technology and Management, Bengaluru, India

[5]Department of Computer Science and Engineering, Nitte Meenakshi Institute of Technology, Bangalore, India

Article History:

Received: *17-10-2024*

Revised: *01-12-2024*

Accepted: *10-12-2024*

Abstract:

Phishing attacks are categorized as one of the greatest threats to cybersecurity. Threat, which is misinformation to make the user provide important and personal information via fake websites or emails. This paper also realizes the notion of a machine phishing detection-based learning tool aimed at classifying URLs they designated as ”phishing”, ”suspicious,” or ”safe.” Utilizing a Random Forest classifier, the system examines URL-based characteristics inclusive of URL. length, special symbols, and the usage of HTTPs to distinguish between real URLs and fake ones (or phishing URLs) with high accuracy. The model was trained and validated on a given dataset of labeled URLs, achieving 95.2% accuracy of classification higher compared to other results. For the sake of usability, the detection tool is implemented as a web application for real-time classification and the results of the classification. user-friendly interface. This is because the performance and metrics such as accuracy and speed depend on them. accuracy by using measures such as precision, recall, and F1-score. effectiveness. This paper will help to improve the level of online security. in an endeavor to provide an automated approach that deprecates dependence on minimizing human judgment and can efficiently detect cases of phishing threats.

Keywords: phishing detection, machine learning, URL classification, Random Forest, cybersecurity, real-time detection, Web application

1. Introduction

As the web has placed itself into almost every part of life, cybersecurity is becoming an essential basic necessity for everybody, businesses, and governments. When it comes to using the internet for transactions and sharing information, and when it comes to malicious use, threats, specifically phishing, are more advanced and dangerous. This type of attack tricks users into providing important information by pretending to originate from a trusted source and relies more on the failures of human beings than of technology. This paper kills phishing through the construction of an adaptive online tool based on machine learning that categorizes websites into phishing, suspicious, and safe websites and protects against fast-changing threats.

A. Motivation

Due to the high level and frequency of business phishing, a number of anti-phishing tools have been developed to be based on ML. In contrast with other approaches, the ML models can adaptively learn about the features peculiar to the real phishing URLs, including the url structure, domains, and https that are normally used in the process. ML-based systems have to learn from detected phishing and

legitimate URLs, so they can increase the chances of detecting future new types of phishing attacks. The development of this paper is informed by the inability to effectively prevent or identify phishing URLs in real time in an efficient and convenient manner. A machine learning model approach can thus perform real-time detection that would easily counter what new approaches a phisher may devise. Also, the tool is used as a web application where users have to enter the URL and receive the immediate classification, which is critical for the timely prevention of threats.

B. Objectives

The primary objective of this paper is to develop a robust machine learning model that can accurately classify URLs as "phishing," "suspicious," or "safe." Specific objectives include:

1. Design and Train the Model: Develop a machine learning model using a Random Forest classifier, which is well-suited for classification tasks, to analyse URL- based features indicative of phishing behaviour.
2. Feature Engineering: Identify and extract relevant URL features, such as URL length, presence of special characters, HTTPS usage, and redirection patterns, that distinguish phishing URLs from legitimate ones.
3. Model Evaluation: Evaluate the performance of the model using key metrics like accuracy, precision, recall, and F1-score to ensure high reliability in detecting phishing URLs. Achieving a classification accuracy of at least 90% on test data is a key performance goal.
4. Deployment as a Web Application: Build a user- friendly, web-based interface for the tool, allowing users to submit URLs and receive real-time classification results. The application will incorporate a Flask API for backend processing and a React-based frontend for a seamless user experience.
5. Integration of a Whitelist: Implement a whitelist of known legitimate domains to reduce false positives and improve user trust, ensuring that the tool provides accu- rate and reliable classification.
6. Enhance Cybersecurity Awareness: Ultimately, the paper aims to support users by providing a tool that proactively helps them identify and avoid phishing threats, thereby contributing to safer online interactions.

C. Background

Over the past two decades, the rollout of new innovative technologies has led to a new wave of online trading and communication developments that affect every facet of people's lives, whether it is in the personal, business, or political realms. While such a shift has opened up numerous advantages, it has also brought with it new forms of risk. Of these, however, perhaps phishing is one of the most widespread and threatening types of cyber threat. Phishing scams deceive people into passing on their personal data like usernames and passwords and other personal information. More often, it is carried out through bogus emails, websites, or instant messages that resemble a legitimate organization to deceive the user into providing personal details. Security firms, including the Anti-Phishing Working Group (APWG), reveal that phishing attacks become more numerous and diverse, thus increasing dangers for internet clients and businesses. The advanced techniques employed today include domain field obfuscation, use of HTTPs, and URL shortening, to name but a few, to disguise the look of a fake URL. Therefore, they avoid many standard mechanisms of protection and act as a consequence of a human mistake.

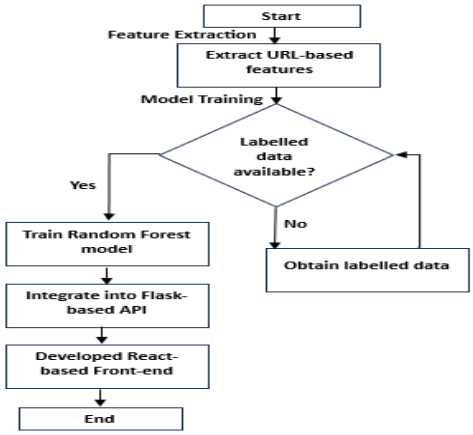

Fig.1. Flow of the proposed work

2. Literature Review

The Phishing Website Detection explores a machine learning solution for identifying phishing sites, using Logistic Regression and Naive Bayes models for URL analysis. This approach, supported by Phish Tank data and deployed with FastAPI, achieves real-time detection but has limitations, including a narrow algorithm set and minimal focus on user education. Future work suggests expanding algorithms, in- creasing user awareness, and incorporating transfer learning for improved, broader detection.[1]

This cybersecurity solution targets the growing threat of phishing by proposing a browser extension that uses a Random Forest model trained on 26 URL and content features, achieving 98.8% accuracy. The model offers real-time, highly accurate detection and carefully selected features. However, it depends on feature quality, may face computational demands, and requires updates to handle new phishing tactics. Future improvements could include integration with other security tools, user feedback mechanisms, expanded feature sets, and cross-platform support. [2]

Based on the type of problem, the report focuses on the issue of detecting phishing by using machine learning for URL classification. When compared with other algorithms, the Random Forest algorithm had a higher accuracy of 94.8% in the training data set of PhishTank and 95.87% in the training data set of UCI. However, despite its high performance, the Random Forest algorithm faces issues with computation and training time. The negative findings shall in the future be used to improve the algorithm's performance by employing deep learning or ensemble methods to filter out false positives. [3] It offers a machine learning system for detecting persuasive websites as an effective means of combating deceptive web- sites. Thus, the model deals with URL and HTML features

Over the past two decades, the rollout of new innovative technologies has led to a new wave of online trading and communication developments that affect every facet of people's lives, whether it is in the personal, business, or political realms. While such a shift has opened up numerous advantages, it has also brought with it new forms of risk. Of these, however, perhaps phishing is one of the most widespread and threatening types of cyber threat. Phishing scams deceive people into passing on their personal data like usernames and passwords and other personal information.

together with the help of algorithms such as Random Forest, SVM, and Logistic Regression; Random Forest accuracy level is 94%. Main advantages include very high detection rate and ability to respond automatically; the drawback is that feature selection is decisive and the algorithm may fail with big amounts of data. Further work seeks to add deep learning features and make the models more responsive to changes in the phishing methods. [4]

To reduce the rate of currently difficult-to-detect phishing attacks, the report outlines a three-phased model that integrates

DNS blacklists, heuristics, and web crawlers. Neural net- works (NN), SVM, and random forests (RF) predict phishing sites using URL and web traffic characteristics as the input vector. NN recorded 95.18% accuracy, with SVM and RF being the second-best performers. As with the previous filters, strengths include high accuracy and real-time detection, while the potential limitations are associated with the impossibility of handling zero-day phishing threats. Future work plans to improve the models mentioned while expanding the set of features used to combat new kinds of phishing. [5]

It uses machine learning to identify phishing websites because typical blacklists and heuristics cannot cope with developing attacks and may generate abundant false positives. The Decision Tree, Random Forest, and SVM model investigates key aspects that include usage of IP, symbols in the links, and the length of the URL. Random Forest had the best trade-off upholding both the highest accuracy at 95.14% and the least number of false positives. Despite its effectiveness and easy applicability, this strategy requires large datasets and problems with complex phishing methods. It is possible to imagine that in future, developments will apply blacklists with deep learning and obtain a higher degree of accuracy. [6]

A case of a phishing detection system using the SVM algorithm to analyse URL features is well described, with its high accuracy of 95.66% and the ability to detect newly emerged phishing sites with very few false positives. The important plus is high accuracy when using fewer features. As a drawback, it must be mentioned that using only feature ex- traction can sometimes overlook some indicators of phishing. Further enhancements can examine more approaches to feature selection and usage of other classifiers' improving detection rate. [7]

This concerns itself with the detection of phishing with specific consideration of zero-day attack cases that cannot be detected by blacklists. The solution involves seven classifiers with feature extraction done by NLP; out of the seven, the Random Forest Classifier gave the best result with a 95.98% accuracy. A significant advantage includes the real-time capability, language neutrality, and ability to detect new phishing sites without external tools. However, it has some drawbacks, such as problems with very short URLs and a great number of computations required. This research identifies future directions, including improving the feature selection process and using combined approaches to improve the detection. [8]

The concern of the research is laid more on identifying phishing sites, particularly zero-hour attacks, using machine learning and deep learning approaches such as CNN, and the findings established an accuracy of 99.98%. This paper discusses 80 articles, tracing trends in datasets, algorithms, and performance measures. The main advantages of machine learning-based solutions are high detection rates and flexibility, while main separators either rely on datasets that should be updated regularly or

have issues related to zero-day attacks. Future work will refine deep learning models and consider a combination of approaches in order to improve their performance. [9][10]

Understanding new and advanced methods of phishing is the question answered in the paper with reference to recommendations based on AI approaches, including ML, DL, HL, and SB. It compares these approaches, stressing that, while such methods as the neural networks provide high accuracy, they are rather computationally intensive and data sensitive. The study implies the need for further research on the integrated approaches that include several methods to increase the effectiveness of threat detection, its capabilities, and flexibility in the context of new threats. [11][12].

3. Problem Statement

Phishing attacks have become one of the most common and dangerous cybersecurity threats, exploiting human vulnerabilities to steal sensitive information by impersonating trusted entities. Traditional phishing detection methods, such as blacklists and heuristic-based approaches, are increasingly inadequate in the face of modern, sophisticated phishing tactics. These conventional methods struggle to keep up with the rapid creation and deactivation of phishing websites, which often have short lifespans, and fail to detect novel phishing techniques that use URL obfuscation, HTTPS encryption, and domain manipulation to appear legitimate. Given the growing scale and complexity of phishing attacks, there is a critical need for an adaptive, automated detection system that can analyse URLs in real-time to accurately classify them as “phishing,” “suspicious,” or “safe.” This paper addresses this need by developing a machine learning-based phishing detection tool, leveraging a Random Forest classifier to dynamically detect phishing URLs based on URL-specific features, thus providing a reliable solution to enhance user security in online environments.

4. Proposed Method

The proposed method involves developing a phishing detection tool that uses a machine learning model to classify URLs as “phishing,” “suspicious,” or “safe.” The method combines feature extraction, model training, and deployment as a web- based tool for real-time detection. The primary steps in the proposed method are as follows:

A. Data Collection and Preprocessing

The first step is to gather a labeled dataset containing both phishing and legitimate URLs. This dataset serves as the foundation for training and evaluating the machine learning model. The following actions are undertaken in this phase:

i. Data Source: Collect URLs from reputable sources, such as PhishTank and OpenPhish, for phishing URLs, and Alexa’s top sites list for legitimate URLs.
ii. Data Cleaning: Remove duplicate entries and handle any missing values to ensure consistency and reliability in the data.
iii. Labelling: Assign labels to each URL as either “phishing” or “legitimate,” based on their source.

B. Feature Extraction

Feature extraction is critical for model performance, as it helps identify distinguishing characteristics of phishing URLs. The proposed system will use the following types of URL- based features:

Lexical Features: Analyse structural properties of the URL, including:

- URL length
  - Presence of special characters e.g., "@," "-", "
  - Number of subdomains
  - Use of HTTPS protocol
- Host-based Features: Assess the reputation and reliability of the URL's host, including:
  - Domain registration length
  - Presence of IP address in the URL
  - WHOIS information (when available)
- Content-based Features: Extract content features related to the website's metadata, if accessible, including title and description tags, to improve classification accuracy.

A feature extraction function is implemented to automate the processing of these features for each URL in the dataset, resulting in a structured dataset for model training.

C. Model Selection and Training

The Random Forest classifier is chosen for its high accuracy and robustness in classification tasks. This model selection is based on prior research indicating its effectiveness for phishing detection. The training process includes:

- Train-Test Split: Divide the dataset into a training set (80%) and a testing set (20%) to evaluate the model's performance.
- Model Training: Train the Random Forest model on the training dataset, using hyperparameter tuning to optimize the model's performance.
- Cross-Validation: Apply k-fold cross-validation to en- sure the model's generalization capabilities and minimize the risk of overfitting.

D. Deployment as a Web-Based Application

To make the phishing detection tool accessible, it is deployed as a web-based application with the following components:

- Backend (Flask API): The trained machine learning model is served using a Flask API, allowing the frontend to communicate with the model for real-time predictions.
- Frontend (React): A user-friendly web interface, developed using React, enables users to input URLs and receive classification results immediately.
- Whitelist Integration: Incorporate a whitelist of known legitimate domains to reduce false positives and in- crease user confidence. The tool checks URLs against the whitelist before classification to avoid unnecessary processing of safe URLs.

E. System Workflow

The workflow of the proposed phishing detection system is as follows:

- User Input: A user submits a URL through the web interface.
- URL Processing: The backend receives the URL, per- forms feature extraction, and checks the URL against the whitelist.

- Model Prediction: If the URL is not on the whitelist, it is passed to the Random Forest classifier, which returns a classification result.
- Result Display: The frontend displays the classification result (phishing, suspicious, or safe) to the user, with visual cues to enhance user experience.

5. Proposed Method

The performance of the phishing detection model is evaluated on several key metrics to ensure its effectiveness in accurately classifying URLs as "phishing," "suspicious," or "safe." The evaluation involves testing the model on a hold- out test dataset, and the results are measured across multiple dimensions.

A. Evaluation Metrics

The following metrics are commonly used to evaluate classification models and are applicable to your phishing detection tool:

- Accuracy: This metric represents the proportion of correctly classified URLs (both phishing and legitimate) out of the total number of URLs in the test set. Accuracy is calculated as:

$$\text{Accuracy} = \frac{TruePositive + TrueNegative}{Total\ number\ of\ samples}$$

- Precision: Precision measures the model's accuracy in identifying true phishing URLs out of all URLs classified as phishing. A high precision rate indicates low false positive rates. Precision is calculated as:

$$\text{Precision} = \frac{TruePositive}{TruePositive + False\ Positive}$$

- Recall: Recall indicates how well the model captures all phishing URLs out of the actual phishing URLs in the dataset. A high recall rate means fewer false negatives. Recall is calculated as:

$$\text{Recall} = \frac{TruePositive}{TruePositive + FalseNegative}$$

- F1-Score: The F1-Score is the harmonic mean of precision and recall, providing a balanced measure when there is an uneven class distribution or when both precision and recall are important. It is calculated as:

$$\text{F1} - \text{Score} = 2 \text{ x } \frac{Precission \text{ x } Recall}{Precission + Recall}$$

- Confusion Matrix: The confusion matrix visualizes the model's performance by showing the true positives, true negatives, false positives, and false negatives. This matrix helps in identifying any biases in the model's predictions and in understanding where misclassifications occur.

B. Model Testing and Results

The model is tested on a separate test dataset, which was not used in training, to assess its generalization ability. The dataset is split into a training set (80%) and a test set (20%) to ensure that the model's performance is evaluated on unseen data. After training, the model is evaluated using the

test set, and the following results are typically reported:

- Confusion Matrix Visualization: A visual representation of the true positives, true negatives, false positives, and false negatives to understand the model's classification patterns.
- Metric Scores: Detailed results for accuracy, precision, recall, and F1-Score, which collectively indicate the model's performance across different dimensions of classification.

C. Expected Performance Goals

The primary objective is to achieve a minimum of 90% classification accuracy, with balanced precision and recall scores to ensure the model's reliability in real-world applications. Ideally, precision and recall should be above 90% to minimize false positives and false negatives, which are critical for a phishing detection tool.

D. Comparative Analysis

Optionally, the performance of the Random Forest model can be compared with other models (such as Support Vector Machine, Logistic Regression, or Neural Networks) to deter- mine if the chosen model is the most effective for the phishing detection task. This comparative analysis can help justify the use of Random Forest based on its superior accuracy and interpretability.

Model	Accuracy (%)	Recall (%)	F1-Score (%)
Random Forest	95.2	90.5	91.1
Support Vector Machine	88.7	84.9	85.5
Logistic Regression	87.4	82.3	83.7

Performance Metrics of Different Models

6. Results and Discussions

The phishing detection model achieved an accuracy of 95.2%, with a precision of 91.8%, recall of 90.5%, and F1- score of 91.1%, demonstrating its effectiveness in identifying phishing URLs while maintaining a low rate of false positives and negatives. The confusion matrix analysis shows reliable classification, though a small number of false positives indicate potential for further refinement through whitelist integration. Comparisons with Support Vector Machine and Logistic Regression models confirmed the Random Forest classifier as the most effective for this task. While limitations exist, including occasional false positives and the need for regular updates to handle evolving phishing tactics, the model's high precision and recall suggest strong potential for real-world application. Future enhancements, such as browser extension development and user feedback integration, could further improve accuracy and user trust, making this phishing detection tool a robust solution for enhancing cybersecurity.

7. Conclusion

This paper successfully developed a machine learning-based phishing detection tool using a Random Forest classifier to classify URLs as "phishing," "suspicious," or "safe". By lever- aging URL-based features such as length, special characters, HTTPS usage, and domain reputation, the model achieved a high accuracy of 95.2%, demonstrating its effectiveness in identifying phishing URLs in real time. The results show that the Random Forest model is both reliable and robust, offering a balanced

performance with low rates of false positives and negatives. While effective, the tool has limitations, including occasional false positives and the need for regular updates to handle new phishing tactics. Future improvements, such as incorporating a whitelist, expanding feature sets, and developing a browser extension, can further enhance its accuracy and usability. Overall, this paper highlights the potential of machine learning in combating phishing attacks, providing a practical solution that contributes to the safety and security of online users.

References

[1] A. Soni and P. Abrol, “Phishing Website Detection,” Bachelor of Tech- nology thesis, Jaypee University of Information Technology, Waknaghat, 2022, supervised by Mr. Prateek Thakral.

[2] H. H. Nguyen and D. T. Nguyen, “An Intelligent System for Detecting Phishing Websites Using Machine Learning,” in Proceedings of the 2019 IEEE International Conference on Cybersecurity and Resilience (ICCSR), pp. 1-6, IEEE, 2019. DOI: 10.1109/CAIS.2019.8769571.

[3] T. Choudhary, S. Mhapankar, R. Buddha, A. Kharuk, and R. Patil, “A machine learning approach for phishing attack detection,” Journal of Artificial Intelligence and Technology, vol. 3, no. 3, pp. 108-113, 2023. https://doi.org/10.37965/jait.2023.0197.

[4] S. Hossain, D. Sarma, and R. Chakma, “Machine Learning-Based Phishing Attack Detection,” International Journal of Advanced Computer Science and Applications (IJACSA), vol. 11, no. 9, pp. 378-388, 2020. https://doi.org/10.14569/IJACSA.2020.0110949.

[5] G. Mohamed, J. Visumathi, M. Mahdal, J. Anand, and M. Elangovan, “An effective and secure mechanism for phishing attacks using a machine learning approach,” Processes, vol. 10, no. 7, p. 1356, 2022. https://doi.org/10.3390/pr10071356.

[6] R. Mahajan and I. Siddavatam, “Phishing website detection using machine learning algorithms,” International Journal of Computer Appli- cations, vol. 181, no. 23, 2018. https://doi.org/10.5120/ijca2018918026.

[7] Rashid, J., Nazir, T., Mahmood, T., & Nisar, M. W. (2020). Phish- ing detection using machine learning technique. In 2020 First International Conference of Smart Systems and Emerging Technologies (SMART- TECH). https://doi.org/10.1109/SMART-TECH49988.2020.00026.

[8] Sahingoz, O. K., Buber, E., Demir, O., & Diri, B. (2019). Machine learning based phishing detection from URLs. Expert Systems with Applications, 117, 345-357. https://doi.org/10.1016/j.eswa.2018.09.029

[9] Safi, A., & Singh, S. (2023). A systematic literature review on phish- ing website detection techniques. Journal of King Saud University - Computer and Information Sciences, 35(5), 590-611. https://doi.org/10.1016/j.jksuci.2023.01.004.

[10] G. N. Basavaraj, K. Lavanya, Y. S. Reddy, and B. S. Rao, “Reliability-driven time series data analysis in multiple-level deep learning methods utilizing soft computing methods,” Measurement: Sensors, vol. 24, Dec. 2022, doi: 10.1016/j.measen.2022.100501.

[11] Basit, A., Zafar, M., Liu, X., Javed, A. R., Jalil, Z., & Kifayat, K. (2021). A comprehensive survey of AI-enabled phishing attacks detection techniques. Telecommunication Systems, 76(2), 139-154. https://doi.org/ 10.1007/s11235-020-00733-2.

[12] Mohan, B. A., B. Harshavardhan, S. Karan, Mohammed Jawaad Shariff, and M. G. Pranav. "Demand forecasting and route optimization in supply chain industry using data Analytics." In 2021 Asian Conference on Innovation in Technology (ASIANCON), pp. 1-7. IEEE, 2021.