© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1716247

Phishing Website Detection Using Machine Learning

MUGUNTHAN KENNEDY K[1], DR. S. SRIDEVI[2], HARIESH KUMAR P[3], NANTHAKUMAR R[4]
[1,2,3,4]*Computer Science and Engineering Department, Velammal Engineering College, Chennai,India*

Abstract- ***Phishing attacks have been one of the biggest cybersecurity concerns, where attackers develop fraudulent websites that resemble legitimate online services to steal critical user information such as login credentials, bank account information, and other sensitive information. The conventional anti-phishing detection systems, such as blacklist-based systems, rule-based systems, etc., are found to be less effective in handling the detection of new phishing sites that emerge frequently in the world wide web. To overcome the limitations of the conventional systems, the authors of the current research propose a supervised machine learning-based anti-phishing detection system using a set of discriminative features extracted from the address bar, domain-based features, and other components of the webpage such as HTML and JavaScript code. A set of machine learning and deep learning-based classifiers have been trained and tested using a balanced set of legitimate and phishing URLs to evaluate the performance of the proposed system. The experimental results show that the proposed system achieves better performance in terms of accuracy, precision, and recall using ensemble-based models such as XGBoost in comparison to other state-of-the-art models.***

Index Terms- ***Phishing Detection, Machine Learning, Cyber Security, URL Features, Ensemble Learning.***

I. INTRODUCTION

In addition, the rapid expansion of the internet and online services has greatly increased the risks of various cyber attacks, among which phishing has become one of the most common and dangerous threats. Basically, phishing is a kind of social engineering attack where attackers try to trick users by launching fake websites that resemble genuine websites. These fake websites aim at tricking users into providing their personal information such as usernames, passwords, credit card numbers, etc.

Conventional detection of phishing attacks usually depends on blacklist-based detection, heuristic detection, and signature-based detection. Blacklist-based detection usually involves maintaining a database of known phishing sites, which blocks access if the URL is found on the database of known phishing sites. This approach, however, is found to be ineffective for new sites that may not be included on the database of known phishing sites. Heuristic-based detection usually depends on known patterns of phishing attacks, which can be easily evaded by attackers who keep on modifying their attack strategies.

As a result of the limitations of traditional methods, machine learning methods have attracted significant interest in the field of phishing detection. Machine learning algorithms have the potential to learn from large datasets of phishing and legitimate websites and make accurate predictions on new URLs based on their characteristics. Machine learning algorithms have the potential to identify various features and learn from them to recognize the differences between phishing and legitimate websites.

In this study, a machine learning-based phishing website detection system has been proposed. The proposed system has the potential to learn various discriminative features from the address bar, domain, and webpage content. The proposed system has the potential to train various machine learning and deep learning models, including Decision Tree, Random Forest, Support Vector Machine (SVM), XGBoost, Multilayer Perceptron, and Autoencoder Neural Networks, to determine the effectiveness of these models in detecting phishing websites.

The primary aim of this study was to propose an intelligent and automated phishing website detection system, which has the potential to accurately detect phishing websites and reduce false alarms. The proposed phishing detection system has the potential to be an effective tool in modern-day cybersecurity applications.

© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1716247

II. PHISHING TECHNIQUES

The techniques used in phishing attacks are called phishing techniques. Phishing techniques are ways in which attackers trick people into giving their sensitive information such as usernames, passwords, bank account details, etc. Phishing techniques are created in a way that people think they are using a genuine website or service. Phishing techniques have been evolving over time and have become more advanced. Here are some of the most common phishing techniques used by attackers.

A. Link Manipulation

The most common technique used in phishing attacks link manipulation. In link manipulation, attackers use a technique in which attackers send a link to a user that seems to be a genuine website but actually is a phishing website. Attackers use a technique in which attackers send a link in a text format that is not easily visible to a user. In some cases, attackers use a technique in which attackers use a different spelling of a website's name or use a different character in a website's name.

B. Website Forgery

Website forgery occurs when an attacker creates animitation of an original legitimate website with the aim of deceiving users into disclosing confidential information. The attackers create an imitation of legitimate websites such as banking websites, email services, or online stores. The attackers then use the imitation website to collect confidential information from users. The users may find it difficult to differentiate the legitimate website from the imitation website since the imitation website resembles the legitimate website.

C. Hidden Redirects

Hidden redirects are used for silent redirects from an original legitimate webpage to a phishing website without the knowledge of the users. The attackers use malicious scripts or redirection code on the web pages, advertisements, or emails. When the users click on the link, they are redirected to a malicious website that aims at obtaining the users' information. This technique enables the attackers to evade basic detection systems.

D. Image-Based Phishing

This is a type of phishing where the attacker utilizes images rather than text to display important information on the webpage. The attacker has a higher probability of avoiding detection since most phishing detectors are programmed to analyze the content of the webpage, and images are not included in the content analysis process.

E. Email Spoofing

This is another type of phishing where the attacker sends emails that seem to be from trusted sources, such as banks, the government, or well-known companies and organizations. The email contains messages that are urgent and require the user to update their information or click on a suspicious link, among other things. The attacker has a higher probability of tricking the user into falling into the phishing trap because the email is from a trusted source.

In conclusion, it is important to note that all the above phishing techniques are based on deception and are used to take advantage of human trust.

III. PHISHING DETECTION APPROACH AN OVERVIEW

Considering the rapid increase in phishing attacks, various detection mechanisms have been developed to protect internet users from malicious websites. The detection mechanisms have been developed to identify phishing sites before users access them and provide their sensitive information. There exist two categories of phishing detection mechanisms: traditional security approaches and machine learning-based approaches.

A. Blacklist-Based Detection

Blacklist-based detection is one of the widely used methods to detect phishing sites. In this method, a database of phishing sites' URLs is maintained by security organizations. When a user tries to access a website, the URL of the website is compared with the URLs in the blacklist database. If the URL matches the phishing sites' database, access to the website is blocked, and a warning message is displayed to the user.

Even though this detection method is effective in detecting previously identified phishing sites, this method has various limitations. Blacklist detection mechanisms fail to identify newly created phishing sites, which have not been added to the database. The phishing sites created by attackers for a short period of time will be closed before they are identified and added to the blacklist database.

B. Heuristic-Based Detection

Heuristic-based detection mechanisms analyze the characteristics of URLs and webpages to identify if there are any signs that indicate a webpage is suspicious or a phishing site. Heuristic-based systems are defined by specific rules that identify abnormal patterns usually found in phishing webpages. The abnormal patterns include excessively long URLs, the use of IP addresses, and the presence of suspicious characters in the URL.

Even though heuristic-based systems are capable of detecting unknown phishing webpages, the systems are limited in that attackers can easily evade them by changing their phishing strategies. This makes the effectiveness of heuristic-based systems limited.

C. Signature-Based Detection

Signature-based systems detect phishing webpages by analyzing patterns or signatures in the content, HTML, or JavaScript code found in a webpage. If a webpage contains a signature or pattern similar to that found in a known phishing webpage, then it is classified as a phishing site.

This type of system is effective in detecting known phishing patterns, but it is limited in that it cannot detect new or altered phishing attacks. Since phishing attacks are constantly changing, signature-based systems must be updated regularly to be effective in detecting phishing attacks.

D. Machine Learning-Based Detection

Machine learning-based detection has also been widely studied in the past few years because of the ability of the system to learn patterns from the given data. The system can analyze different features of the URLs, domains, and structure of the webpage to classify the website as a legitimate one or a phishing one.

Decision Trees, Random Forest, Support Vector Machines, Neural Networks, and XGBoost are some of the machine learning algorithms that can be used in the detection of phishing sites. The main advantage of using machine learning-based detection is that the system can learn from the past data and detect new phishing sites that were not in the training set of the system, as opposed to the traditional detection mechanisms that can only detect known types of attacks.

Machine learning-based detection systems can be said to be intelligent in the sense that the system can learn patterns from the given data and can be much better than the traditional detection mechanisms in the detection of phishing sites.

IV. MACHINE LEARING APPROACH

The use of machine learning techniques is an effective solution in detecting phishing websites because of their ability to learn from large data sets and automatically classify unknown URLs. In machine learning-based phishing detection systems, a binary classification is performed on a website, determining whether a website is a phishing site or a legitimate site. Machine learning models use various features extracted from URLs and content to learn patterns that can differentiate between phishing and legitimate websites.

A. Data Collection

The data set used for training and testing machine learning models consists of phishing URLs and legitimate URLs. Phishing URLs are obtained from publicly available sources containing phishing URLs, whereas legitimate URLs are obtained from trusted data sets containing genuine URLs of websites. The data set is balanced in terms of phishing and legitimate websites.

B. Feature Extraction

Feature extraction is an important step in phishing website detection. In this paper, a total of 17 discriminative features are extracted from various parts of the website. The features can be summarized into three categories:

- Address Bar-Based Features: These features examine the URL structure of a website. The features include URL length, use of special characters in URLs, use of IP addresses instead of domain names, and use of suspicious prefixes and suffixes.
- Domain-Based Features: These features examine various pieces of information related to a website's domain.
- HTML and JavaScript-Based Features: These features examine various pieces of information related to a website's behavior and structure.

These features enable machine learning algorithms to recognize hidden patterns that are usually associated with phishing websites.

C. Model Training

After feature extraction, a division of the data into a training and a testing set is made, normally with a proportion of 80:20. The training data set is used to train the models, and the testing data set is used to test these models.

Various machine learning and deep learning algorithms are implemented and compared in this study. Some of these algorithms include

- Decision Tree
- Random Forest
- Support Vector Machine
- XGBoost
- Multilayer Perceptron
- Autoencoder Neural Network

Each model learns and tries to classify these websites accurately as either phishing or normal.

D. Model Evaluation

After training these models, they are evaluated using standard metrics such as accuracy, precision, recall, and F1-score. These metrics can be used to determine how effectively these models can detect these phishing websites without resulting in false positives or false negatives.

Among these models, ensemble-based models such as XGBoost perform better than other models because they can effectively combine multiple models to create a single model.

As can be seen, the proposed machine learning model provides an intelligent and automatic framework for detecting these phishing websites and can effectively identify both known and unknown types of phishing websites.

V. DATASET DESCRIPTION

The success of a machine learning-based phishing detection system is greatly dependent on the quality and reliability of the dataset used for training and testing. In this paper, a dataset containing both phishing and legitimate URLs is used for training the machine learning models. The dataset is created in a manner that provides a balanced representation of phishing and legitimate websites.

The phishing URLs are obtained from publicly available repositories containing updated information on newly detected phishing websites. These repositories contain verified phishing URLs that are used for various research purposes. Similarly, the legitimate URLs are obtained from sources containing verified and safe websites. Combining both types of URLs provides a representation of both malicious and legitimate web behavior.

The dataset used in this paper contains 10,000 URLs in total, of which 5,000 URLs are phishing URLs and 5,000 URLs are legitimate URLs. The phishing URLs are obtained from a publicly available phishing URL repository called PhishTank. PhishTank is a repository containing a list of updated phishing websites reported and verified by various users in the online community. Similarly, the legitimate URLs are obtained from the University of New Brunswick's publicly available cybersecurity datasets containing various types of web traffic data.

From these URLs, a list of 17 significant features is derived to describe the characteristics of each website. These features fall into three broad categories:

1. Address Bar-Based Features – This category of features focuses on analyzing the characteristics of the URL, such as the length

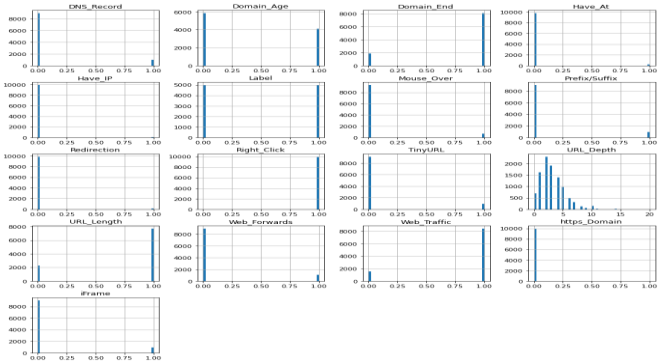
© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1716247

of the URL, the presence of special characters, and the use of IP addresses instead of domain names.

2. Domain-Based Features – This category of features focuses on analyzing the characteristics of the domain, such as domain age and website popularity.
3. HTML and JavaScript-Based Features – This category of features focuses on analyzing the characteristics of the webpage, such as embedded scripts and abnormal webpage elements.

A. Address Bar-Based Features (9 Features)

1. Using IP Address
   Phishing sites often make use of their IP address rather than their domain name to hide their ownership.
2. URL Length
   Phishing sites often make use of long URLs to hide their malicious parts.
3. Shortening Service
   Phishing sites often make use of URL shortening services to hide their malicious parts.
4. Having "@" Symbol
   The presence of the "@" symbol in the URL often redirects the user to another website.
5. Double Slash Redirecting ("//")
   The presence of more than one "//" in the URL, excluding the protocol part, often redirects to another malicious website.
6. Prefix or Suffix in Domain
   Phishing sites often make use of the hyphen prefix in their domain to mimic genuine sites, e.g., secure-paypal.com.
7. Subdomain Level
   Phishing sites often make use of more than one subdomain in their URL.
8. HTTPS Token in Domain
   Phishing sites often make use of the term "https" in their domain to appear more secure.
9. URL Redirecting
   Phishing sites often make excessive use of URL redirecting.

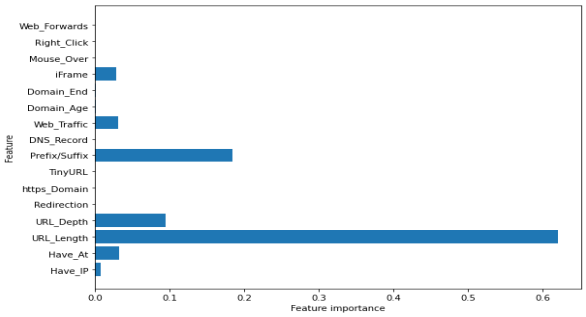


B. Domain-Based Features (4 Features)

1. Domain Registration Length
   Websites often make long-term registrations for their domain, but phishing sites make short-term registrations.
2. Age of Domain
   Websites often make new domain registrations for phishing activities.
3. DNS Record Availability
   Websites often lack DNS records, indicating suspicious and temporary sites.
4. Website Traffic Rank
   The website traffic rank is usually higher for genuine websites compared to phishing websites.

C. HTML and JavaScript-Based Features (4 Features)

1. Using Iframe
   The attackers use invisible iframes to hide their content.
2. Status Bar Customization
   The attackers use JavaScript code to change the status bar of a webpage.



3. Disabling Right Click
   The attackers use a script to disable the right-click function in order to prevent users from checking the webpage,
4. Using Pop-up Windows
   The attackers use pop-up windows to ask for user credentials.

The above 17 features cover all the characteristics of a phishing website in terms of structure, behavior, and domain.

VI. EVALUATION METRICS

In order to assess the performance of the machine learning models, which have been utilized for the detection of phishing websites, certain evaluation metrics have been utilized. The metrics help in assessing the accuracy of the classification of phishing and genuine websites by the models and their overall performance. The commonly utilized metrics for the evaluation of the models' performance in phishing detection are accuracy, precision, recall, and F1 score.

A. Accuracy

Accuracy refers to the rate of correctly classified instances out of the total number of predictions made by the model. Accuracy assesses the rate at which the model correctly classifies phishing and genuine websites.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Where:

- TP (True Positive) – Number of phishing websites correctly classified as phishing.
- TN (True Negative) – Number of legitimate websites correctly classified as legitimate.
- FP (False Positive) – Number of legitimate websites incorrectly classified as phishing.
- FN (False Negative) – Number of phishing websites incorrectly classified as legitimate.

A higher accuracy value indicates better overall performance of the model.

B. Precision

Precision measures the proportion of correctly identified phishing websites among all websites predicted as phishing. It reflects how reliable the phishing predictions of the model are.

$$Precision = \frac{TP}{TP + FP}$$

High precision indicates that the model produces fewer false alarms when identifying phishing websites.

C. Recall

Recall, also known as sensitivity, measures the ability of the model to correctly identify actual phishing websites from the dataset.

$$Recall = \frac{TP}{TP + FN}$$

A higher recall value means that the model is able to detect a larger proportion of phishing websites.

D. F1-Score

The F1-score is the harmonic mean of precision and recall. It provides a balanced measure of the model's performance when both precision and recall are important.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

The F1-score is particularly useful when the dataset is imbalanced or when both false positives and false negatives must be minimized.

E. Training Time and Testing Time

In addition to the accuracy level, the computational efficiency is also taken into account. The training time is defined as the time taken to train the machine learning model on the dataset, and the testing time is defined as the time taken by the model to classify the given URL of

the website. The more accurate and efficient model is preferred in a real-time phishing detection system.

These evaluation metrics give a comprehensive view of the performance of the machine learning model used in phishing website detection systems.

VII. EXPERIMENTAL RESULTS

For this research study, different machine learning and deep learning algorithms were implemented and tested to assess their ability to detect phishing websites. The implemented algorithms were trained using the features extracted from the dataset and tested using other features to assess their classification ability. The dataset was divided into two parts: one for training and testing the implemented algorithms using an 80:20 ratio.

The implemented algorithms used in this research study to detect phishing websites are Decision Tree Algorithm, Random Forest Algorithm, Support Vector Machine (SVM), Multilayer Perceptron (MLP), Autoencoder Neural Network Algorithm, and XGBoost Algorithm.

The implemented algorithms were trained using the same dataset and tested to assess their performance using different metrics such as accuracy, precision, recall, and F1-score.

The experimental results indicate that all implemented models were successful in detecting phishing websites with acceptable accuracy. However, it was observed that each model performed differently based on the algorithm used and its ability to learn complex patterns in the data set. Among all evaluated models, Decision Tree was easy to understand but was more prone to overfitting issues. Support Vector Machine (SVM), on the other hand, showed promising classification ability but needed more computation to be carried out.

Among all ensemble methods used in this research, Random Forest and XGBoost showed better results than other implemented models. Ensemble learning combines weak models to build a strong predictive model to improve classification accuracy and prevent overfitting issues. In particular, it was observed that XGBoost showed the best results compared to all other models in terms of accuracy, with 86.4% accuracy on the test data set.

The better performance of XGBoost can be explained in terms of its capacity to deal with intricate relationships between features and its effective gradient boosting algorithm. In addition, neural network-based models such as Multilayer Perceptron and Autoencoder also achieved good results, proving the feasibility of deep learning techniques in phishing detection problems.

As a conclusion, the experimental results prove that machine learning-based techniques can successfully detect phishing websites. Hence, machine learning techniques have great potential in developing intelligent and automated phishing detection systems.

VIII. CONCLUSION AND FUTURE WORK

The better performance of XGBoost can be explained in terms of its capacity to deal with intricate relationships between features and its effective gradient boosting algorithm. In addition, neural network-based models such as Multilayer Perceptron and Autoencoder also achieved good results, proving the feasibility of deep learning techniques in phishing detection problems.

As a conclusion, the experimental results prove that machine learning-based techniques can successfully detect phishing websites. Hence, machine learning techniques have great potential in developing intelligent and automated phishing detection systems.

The experimental results showed that the ensemble learning techniques, such as XGBoost, outperform the other models in terms of accuracy, precision, and recall. This proves that the proposed machine learning techniques can successfully identify phishing websites by learning from the data and identifying the suspicious attributes of the URL or the webpage.

The proposed system can be improved for future work by using more data for better accuracy of the proposed system. Moreover, the proposed system can be improved using advanced deep learning techniques for better accuracy. Additionally, the proposed system can be used for the development of advanced tools that can provide better security for users while surfing the internet.

The proposed system can provide better accuracy for detecting phishing websites. Therefore, the proposed system can play an important role in strengthening the security of the internet.

REFERENCES

[1] J. Ma, L. K. Saul, S. Savage, and G. M. Voelker, "Beyond blacklists: Learning to detect malicious web sites from suspicious URLs," in Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, 2009, pp. 1245–1254.

[2] N. Abdelhamid, A. Ayesh, and F. Thabtah, "Phishing detection based associative classification data mining," Expert Systems with Applications, vol. 41, no. 13, pp. 5948–5959, 2014.

[3] S. Marchal, G. Armano, T. Grondahl, K. Saari, N. Singh, N. Asokan, and A. Francillon, "Off-the-Hook: An efficient and usable client-side phishing prevention application," IEEE Transactions on Computers, vol. 66, no. 10, pp. 1717–1733, 2016.

[4] K. L. Chiew, E. H. Chang, S. N. Sze, and W. K. Tiong, "Utilisation of website features for phishing detection," Computers & Security, vol. 50, pp. 84–92, 2015.

[5] O. K. Sahingoz, E. Buber, O. Demir, and B. Diri, "Machine learning based phishing detection from URLs," Expert Systems with Applications, vol. 117, pp. 345–357, 2019.

[6] UCI Machine Learning Repository, "Phishing Websites Data Set," University of California, Irvine, 2019. Available: https://archive.ics.uci.edu/ml

[7] H. Le, Q. Pham, D. Sahoo, and S. C. H. Hoi, "URLNet: Learning a URL representation with deep learning for malicious URL detection," in Proceedings of the 25th International World Wide Web Conference, 2018, pp. 93–102.

[8] F. Salahdine and N. Kaabouch, "Social engineering attacks: A survey," Future Internet, vol. 11, no. 4, p. 89, 2020.

[9] R. Routh and D. Kshirsagar, "Phishing website detection using machine learning techniques," International Journal of Computer Applications, vol. 174, no. 12, pp. 1–6, 2021.

[10] M. Aljabri, S. AlGhamdi, K. Almarhabi, and F. Alharbi, "Improved phishing website detection using ensemble machine learning techniques," Applied Sciences, vol. 13, no. 8, p. 4649, 2023.

[11] S. Garera, N. Provos, M. Chew, and A. D. Rubin, "A framework for detection and measurement of phishing attacks," in Proceedings of the 2007 ACM Workshop on Recurring Malcode, Alexandria, VA, USA, 2007, pp. 1–8.

[12] R. Verma and N. Das, "What you tweet matters: A new approach to phishing detection using machine learning," in Proceedings of the 2017 IEEE International Conference on Data Mining Workshops (ICDMW), New Orleans, LA, USA, 2017, pp. 1109–1114.

[13] M. Aburrous, M. A. Hossain, F. Thabtah, and K. Dahal, "Intelligent phishing detection system for e-banking using fuzzy data mining," Expert Systems with Applications, vol. 37, no. 12, pp. 7913–7921, 2010.

[14] Y. Zhang, J. Hong, and L. Cranor, "Cantina: A content-based approach to detecting phishing web sites," in Proceedings of the 16th International World Wide Web Conference, Banff, Canada, 2007, pp. 639–648.

[15] M. Khonji, Y. Iraqi, and A. Jones, "Phishing detection: A literature survey," IEEE Communications Surveys & Tutorials, vol. 15, no. 4, pp. 2091–2121, 2013.